Title
Evaluating Optimal Individualized Treatment Rules

Author
Luedtke, Alexander Ryan

Publication Date
2016

Peer reviewed|Thesis/dissertation

Evaluating Optimal Individualized Treatment Rules

by

Alexander Ryan Luedtke

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biostatistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mark J. van der Laan, Chair
Professor Nicholas P. Jewell
Associate Professor Jennifer Ahern

Spring 2016

Evaluating Optimal Individualized Treatment Rules

Abstract

Evaluating Optimal Individualized Treatment Rules

by

Alexander Ryan Luedtke

Doctor of Philosophy in Biostatistics

University of California, Berkeley

Professor Mark J. van der Laan, Chair

Suppose we observe baseline covariates, a binary indicator of treatment, and an outcome occuring after treatment. An individualized treatment rule (ITR) is a treatment rule which assigns treatments to individuals based on their measured covariates. An optimal ITR is the ITR which maximizes the population mean outcome. The mean outcome of the optimal ITR is referred to as the optimal value. This dissertation considers three inferential challenges related to these parameters in the large semiparametric model that at most places restrictions on the probability of receiving treatment given covariates.

The first is to develop confidence intervals for the optimal value. Constructing valid confidence intervals for this quantity is surprisingly difficult when the stratum specific treatment effect, also called the blip function, is null with positive probability. This null treatment effect seems possible in many studies. While it has been claimed in the literature that no regular and asymptotically linear (RAL) estimator exists in this case, we prove that RAL estimators of the optimal value can exist in a slightly more general setting. We then describe an approach to obtain root-n rate confidence intervals for the optimal value even when regular estimation is not possible. We also provide sufficient conditions under which our estimator is RAL and asymptotically efficient – a necessary condition is of course that regular estimation is possible under the data generating distribution.

We have thus far assumed that treatment is an unlimited resource so that the entire population can be treated if this strategy maximizes the population mean outcome. In the second part of this dissertation, we consider optimal ITRs in settings where the treatment resource is limited so that there is a maximum proportion of the population that can be treated. We give a general closed-form expression for an optimal stochastic ITR in this resource-limited setting, and a closed-form expression for the optimal deterministic ITR under an additional assumption. We also present an estimator of the mean outcome under the optimal stochastic ITR and give conditions under which our estimator is efficient among all RAL estimators.

Both of the first two inferential challenges considered give parametric-rate confidence intervals for finite-dimensional parameters in our large semiparametric model. In the third

part of this dissertation we focus on developing hypothesis tests and confidence sets for infinite-dimensional parameters that one typically estimates using data adaptive techniques. Parametric-rate inference is not typically expected in this setting. Our primary motivating example concerns the blip function, which is closely related to the optimal ITRs in both the resource-unconstrained and constrained settings. For any fixed function, we give valid hypothesis tests that the blip function is equal to this fixed function. These tests can then be inverted to develop a confidence set for the blip function. Surprisingly, the hypothesis test achieves a parametric rate in the sense that it is consistent against local alternatives converging to the data generating distribution at the rate of one divided by the square root of sample size. We prove the validity of this procedure in great generality that applies far beyond this particular inference problem, and reference several other examples to which it applies. The results in this third component of the dissertation have been developed using the theory of higher-order influence functions.

To all those who made the last four years unforgettable.

Contents

Acknowledgments

I would first like to thank my advisor, Mark van der Laan, for the honor of learning from him. His clarity of thought never ceases to amaze me, and his dedication to sharing this with others is truly inspirational. I cannot imagine a better role model for the advisor that I one day hope to be. I am deeply grateful for the four years that he has devoted to training me and for the many intellectually thrilling experiences we have shared.

I would also like to thank Marco Carone for his mentorship, generosity, and patience, and Antoine Chambaz for his warm welcome into the Parisian statistics community.

I am thankful to the Berkeley faculty for their contributions to my education, especially Maya Petersen, Jennifer Ahern, and Alan Hubbard. I am also grateful to have been surrounded by incredible graduate students for the last four years, including Robin Mejia, Jeremy Coyle, Wenjing Zheng, Linh Tran, Sara Moore, Marcus Stoiber, Oleg Sofrygin, Iván Díaz, Erin Ledell, Sam Lendle, Laura Balzer, Lucia Petito, and Kelly Street.

I would like to thank my Mom and Dad for giving me the many opportunities that made attending graduate school possible, and my sisters Jamie and Shanna for their constant support. Finally, I would like to thank Christiane for her love and encouragement, and for our many shared adventures.

The material in the three primary chapters of this dissertation has been published elsewhere and is joint with co-authors. Chapter 2 was co-authored by Mark van der Laan and appears in the *Annals of Statistics* under the title "Statistical Inference for the Mean Outcome Under a Possibly Non-Unique Optimal Treatment Strategy". Chapter 3 was co-authored by Mark van der Laan and is in press at the *International Journal of Biostatistics* under the title "Optimal Individualized Treatments in Resource-Limited Settings". Chapter 4 was co-authored by Marco Carone and Mark van der Laan and is under review at the *Annals of Statistics* under the title "An Omnibus Nonparametric Test of Equality in Distribution for Unknown Functions".

Chapter 1

Introduction

We first give an overview of the problems considered in this dissertation. We then present our data structure and the key notation that is common across chapters.

1.1 Overview

Suppose one wishes to maximize the population mean of some outcome using some binary point treatment, where for each individual clinicians have access to measured baseline covariates. Such a treatment strategy is termed an individualized treatment regime ITR, and the (counterfactual) population mean outcome under an ITR is referred to as the value of the ITR (Neyman, 1990; Rubin, 1974; Robins, 1986; Pearl, 2009). An ITR with maximal value is referred to as an optimal ITR or the optimal rule, and the value of an optimal ITR is referred to as the optimal value. One can show that any optimal ITR assigns treatment to individuals falling in strata in which the stratum specific average treatment effect, also termed the "blip function", is positive and does not assign treatment to individuals for which this quantity is negative. This problem has received much attention in the statistics literature over the last two decades (see, e.g., Murphy, 2003; Robins, 2003, 2004; Chakraborty and Moodie, 2013). Most of this earlier work in ITRs has worked in restricted statistical models in which the blip function is parameterized by a finite-dimensional parameter. In this work, we answer several questions arising in individualized medicine in a nonparametric model which at most places restrictions on the probability of a patient receiving treatment given covariates.

Chapter 2: Inference for the Optimal Value

Suppose one wishes to know the impact of implementing an optimal ITR in the population, i.e. one wishes to know the optimal value. Inference about this quantity can be used to inform practitioners as to what benefits are possible if the delivery of a treatment is optimized in the population at a strata-specific level. Before estimating the optimal value, one typically estimates the optimal rule. Recently, researchers have suggested applying machine learning algorithms to estimate the optimal rules from large classes which cannot be described by a finite-dimensional parameter (see, e.g., Zhang et al., 2012a; Zhao et al., 2012; Luedtke and van der Laan, 2014a).

Inference for the optimal value has been shown to be difficult at exceptional laws, i.e. probability distributions where there exists a stratum of the baseline covariates that occurs with positive probability and for which treatment is neither beneficial nor harmful (Robins, 2004; Robins and Rotnitzky, 2014). Inference is similarly difficult in finite samples if the treatment effect is very small in all strata, even though valid asymptotic estimators exist in this setting (van der Laan and Luedtke, 2014a). Zhang et al. (2012b) considered inference for the optimal value in restricted classes in which the ITRs are indexed by a finite-dimensional vector. At non-exceptional laws, they outlined an argument showing that their estimator is (up to a negligible term) equal to the estimator that estimates the value of the *known* optimal ITR under regularity conditions. The implication is that one can estimate the optimal value and then use the usual sandwich technique to estimate the standard error and develop Wald-type confidence intervals (CIs). Van der Laan and Luedtke (2014a) and van der Laan and Luedtke (2014b) developed inference for the optimal value when the ITR belongs to an unrestricted class. Van der Laan and Luedtke (2014b) provide a proof that the efficient influence curve for the parameter which treats the optimal rule as known is equal to the efficient influence curve of the optimal value at non-exceptional laws. One of the contributions of Chapter 2 is to present a slightly more precise statement of the condition for the pathwise differentiability of the mean outcome under the optimal rule. We will show that this condition is necessary and sufficient.

Restricting inference to non-exceptional laws is limiting as there is often no treatment effect for people in some stratum of baseline covariates. Chakraborty et al. (2014) propose using the m-out-of-n bootstrap, which draws samples of size m patients with replacement from the data set of size n, to obtain inference for the value of an estimated ITR using an inverse probability weighted (IPW) estimator. This yields valid inference when the treatment mechanism is known or is estimated according to a correctly specified parametric model. They also discuss an extension to a double robust estimator. In non-regular problems, this method yields valid inference if $m, n \to \infty$ and $m = o(n)$. The CIs for the value of an estimated regime shrink at a root-m (not root-n) rate. In addition to yielding wide CIs, this approach has the drawback of requiring a choice of the important tuning parameter m, which balances a trade-off between coverage and efficiency. Chakraborty et al. propose using a double bootstrap to select this tuning parameter.

Goldberg et al. (2014) instead consider truncating the criteria to be optimized, i.e. the value under a given rule, so that only individuals with a clinically meaningful treatment effect contribute to the objective function. These authors then propose proceeding with inference for the truncated value at the optimal ITR. For a fixed truncation level, the estimated truncated optimal value minus the true truncated optimal value, multiplied by root-n, converges to a normal limiting distribution. Laber et al. (2014a) propose instead replacing the indicator used to define the value of a ITR with a differentiable function. They discuss situations in which the estimator minus the smoothed value of the estimated ITR, multiplied by root-n, would have a reasonable limit distribution.

Chapter 2 develops CIs for the value of the optimal ITR. This chapter carefully studies why standard approaches fail in these settings, and proposes a novel solution which (i) yields asymptotically valid CIs even if the blip function is zero with positive probability and (ii) is asymptotically efficient when the efficiency bound is well-defined. This is the first and only solution satisfying these criteria presented in the literature to date.

Chapter 3: Individualized Treatments Under Limited Resources

In Chapter 2, we aimed to estimate the value of the optimal ITR. If treatment is even slightly beneficial to all subsets of the population, then this ITR would suggest treating the entire population. There are many realistic situations in which such a treatment strategy, or any strategy that treats a large proportion of the population, is not feasible due to limitations on the total amount of the treatment resource. In a discussion of Murphy (2003), Arjas observed that resource constraints may render optimal ITRs of little practical use when the treatment of interest is a social or educational program, though no solution to the constrained problem was given (Arjas et al., 2003).

The mathematical modeling literature has considered the resource allocation problem to a greater extent. Lasry et al. (2011) developed a model to allocate the annual CDC budget for HIV prevention programs to subpopulations which would benefit most from such an intervention. Tao et al. (2012) consider a mathematical model to optimally allocate screening procedures for sexually tranmitted diseases subject to a cost constraint. Though Tao et al. do not frame the problem as a statistical estimation problem, they end up confronting similar optimization challenges to those that we will face. In particular, they confront the (weakly) NP-hard knapsack problem from the combinatorial optimization literature (Karp, 1972; Korte and Vygen, 2012). We will end up avoiding most of the challenges associated with this problem by primarily focusing on stochastic treatment rules, which will reduce to the easier fractional knapsack problem (Dantzig, 1957; Korte and Vygen, 2012). Stochastic ITRs allow the treatment to rely on some external stochastic mechanism for individuals in a particular stratum of covariates.

Chapter 3 first formulates the statistical estimation problem in this setting, which itself was a new contribution as of the publishing of Luedtke and van der Laan (2015). It then describes how to estimate the value of the optimal resource-constrainted ITR.

Chapter 4: Inference for Infinite-Dimensional Parameters

Previous works have demonstrated how one can obtain consistent point estimates of the blip function in mean-square or in value of the resulting optimal rule estimate (see, e.g., Zhao et al., 2012; Luedtke and van der Laan, 2014a). Despite exhibiting strong point estimation schemes for this function (infinite-dimensional parameter), the statistical community has not yet considered how one might obtain inference for this parameter. Chapter 4 develops a hypothesis test that the blip function is equal any given fixed function. Astoundingly, this yields valid confidence sets for this infinite-dimensional parameter: one merely needs to invert the test to check if a given function belongs to the confidence set. The hypothesis tests presented in this chapter are developed using the maximum mean discrepancy (MMD) parameter first presented in Gretton et al. (2006), which represents a discrepancy measure between distributions. While Gretton et al. (2006) focused on the MMD between the distributions of observed random variables, we apply the MMD to measure the discrepancy between the distribution of the (unknown but estimable) blip function applied to the observed covariates minus a fixed function (also applied to the observed covariates) and compare the distribution of this difference to the point mass at zero. While we use inference about the blip function as the primary motivation in this chapter, we present the results in general for a rich class of infinite-dimensional parameters. Inference is usually extremely difficult for infinite-dimensional parameters when smoothing approaches are used to estimate the unknown function because the resulting estimators tend to be highly irregular.

To develop our approach, we use techniques from the higher-order pathwise differentiability literature (see, e.g., Pfanzagl, 1985; Robins et al., 2008; van der Vaart, 2014; Carone et al., 2014). Despite the elegance of the theory presented by these various authors, it has been unclear whether these higher-order methods are truly useful in infinite-dimensional models since most functionals of interest fail to be even second-order pathwise differentiable in such models. This is especially troublesome in problems in which under the null the first-order derivative of the parameter of interest (in an appropriately defined sense) vanishes, since then there seems to be no theoretical basis for adjusting parameter estimates to recover parametric rate asymptotic behavior. At first glance, the MMD parameter seems to provide one such disappointing example, since its first-order derivative indeed vanishes under the null. The latter fact is a common feature of problems wherein the null value of the parameter is on the boundary of the parameter space. It is also not an entirely surprising phenomenon given that the MMD achieves its minimum of zero under the null hypothesis. Nevertheless, we are able to show that this parameter is indeed second-order pathwise differentiable under the null hypothesis – this is a rare finding in infinite-dimensional models. As such, we can employ techniques from the recent higher-order pathwise differentiability literature to tackle the problem at hand. To the best of our knowledge, this is the first instance in which these techniques are directly used (without any form of approximation) to resolve an open methodological problem.

1.2 Notation and data structure

Let $O = (W, A, Y) \sim P_0 \in \mathcal{M}$, where W represents a vector of covariates, A a binary intervention, and Y a real-valued outcome. The model $\mathcal{M}$ for P_0 is nonparametric, beyond possible knowledge about the probability of treatment given covariates. We observe an independent and identically distributed (i.i.d.) sample $O_1, ..., O_n$ from P_0. Let $\mathcal{W}$, $\mathcal{A}$, $\mathcal{Y}$, and $\mathcal{O}$ denote the support of W, A, Y, and O, respectively. For a distribution P, define the treatment mechanism $g(P)(A|W) \triangleq \Pr_P(A|W)$. We will refer to $g(P_0)$ as g_0 and $g(P)$ as g. For a function f, we will use $E_P[f(O)]$ to denote $\int f(o)dP(o)$. We will also use $E_0[f(O)]$ to denote $E_{P_0}[f(O)]$ and $\Pr_0$ to denote the P_0 probability of an event.

Throughout we make the positivity assumption that $\Pr_0(0 < g_0(1|W) < 1)$ so that $\Psi(P_0)$ is well-defined. Let

$$\begin{aligned}\bar{Q}(P)(A,W) &\triangleq E_P[Y|A,W],\\ \bar{Q}_b(P)(W) &\triangleq \bar{Q}(P)(1,W) - \bar{Q}(P)(0,W).\end{aligned}$$

We will refer to $\bar{Q}_b(P)$ as the blip function for the distribution P. We will denote to the above quantities applied to P_0 as $\bar{Q}_0$ and $\bar{Q}_{b,0}$, respectively. We will often omit the reliance on P altogether when there is only one distribution P under consideration: $\bar{Q}(A,W)$ and $\bar{Q}_b(W)$. We also let $\Psi_d(P) \triangleq E_P\bar{Q}(d(W),W)$ denote the value of the rule d. Under causal assumptions, $\Psi_d(P)$ is equal to the counterfactual mean outcome if, possibly contrary to fact, the rule d were implemented in the population (Pearl, 2009). As the focus of this manuscript is statistical in nature, all of the results we present will hold regardless of the validity of the causal assumptions.

The objectives of Chapters 2 through 4 can now be formulated using this notation. In Chapter 2, we wish to develop confidence intervals for the optimal value, defined as

$$\text{Maximize } \Psi_d(P) \text{ over all } d : \mathcal{W} \to \{0,1\}.$$

In Chapter 3, we wish to develop confidence intervals for the optimal resource-constrained value, defined as

$$\text{Maximize } \Psi_d(P) \text{ subject to } E_{P_0}[d(V)] \leq \kappa,$$

where $\kappa \in (0,1)$ is a resource constraint representing the restriction on the proportion of the population able to receive treatment. The definition of Ψ_d is modified very slightly in Chapter 3 to allow d to be a stochastic rule – we refer the reader to that chapter for further details, though we note that in most cases the maximal value over stochastic rules satisfying the constraint is the same as the maximal value over deterministic rules in the above. In Chapter 4, we wish to test

$$\mathcal{H}_0 : \bar{Q}_{b,0}(W) = f(W) \text{ almost surely}$$

for a given fixed function $f : \mathcal{W} \to \mathbb{R}$ against the complementary alternative. From a confidence set perspective, Chapter 4 aims to develop a set of functions CS_n such that

$$\Pr_0\left(\bar{Q}_{b,0} \in \text{CS}_n\right) \to 1 - \alpha$$

for a user-defined confidence level $1 - \alpha$.

Chapter 2

Inference for the Optimal Value

2.1 Introduction

In this chapter, we develop root-n rate inference for the optimal value under reasonable conditions. Our approach avoids any sort of truncation, and does not require that the estimate of the optimal rule converge to a fixed quantity as the sample size grows. We show that our estimator minus the truth, properly standardized, converges to a standard normal limiting distribution. This allows for the straightforward construction of asymptotically valid CIs for the optimal value. Neither the estimator nor the inference rely on a complicated tuning parameter. We give conditions under which our estimator is asymptotically efficient among all regular and asymptotically linear (RAL) estimators when the optimal value parameter is pathwise differentiable, similar to those we presented in van der Laan and Luedtke (2014a). However, they do not require that one knows that the optimal value parameter is pathwise differentiable from the outset. Implementing the procedure only requires a minor modification to a typical one-step estimator.

We believe the value of the unknown optimal rule is an interesting target of inference because the treatment strategy learned from the given data set is likely to be improved upon as clinicians gain more knowledge, with the treatment strategy given in the population eventually approximating the optimal rule. Additionally, the optimal rule represents an upper bound on what can be hoped for when a treatment is introduced. Nonetheless, as we and others have argued in the references above, the value of the estimated rule is also an interesting target of inference (Chakraborty et al., 2014; Laber et al., 2014a; van der Laan and Luedtke, 2014b,a). Thus, although our focus is on estimating the optimal value, the confidence interval presented in this chapter provides proper coverage for the data adaptive parameter which gives the value of the rule estimated from the entire data set under reasonable conditions. We omit these conditions here for brevity, but they can be found in Luedtke and van der Laan (2016).

As in the rest of this dissertation, we focus on the single time point setting in this chapter. We refer the reader to the appendix of the full published paper for the extension to the multiple time point setting Luedtke and van der Laan (2016).

We now give a brief outline of the chapter. Section 2.2 formulates the statistical problem of interest. Section 2.3 gives necessary and sufficient conditions for the pathwise differentiability of the optimal value. Section 2.4 outlines the challenge of obtaining inference at exceptional laws and gives a thought experiment that motivates our procedure for estimating the optimal value. Section 2.5 presents an estimator for the optimal value. This estimator represents a slight modification to a recently presented online one-step estimator for pathwise differentiable parameters. Section 2.6 discusses computationally efficient implementations of our proposed procedure. Section 2.7 discusses each condition of the key result presented in Section 2.5. Section 2.8 describes our simulations. Section 2.9 gives our simulation results. Section 2.10 closes with a summary and some directions for future work.

All proofs can be found in Section 2.11.

2.2 Problem formulation

Let $\Psi : \mathcal{M} \to \mathbb{R}$ be defined by $\Psi(P) \triangleq \Psi_{d(P)}(P)$, where

$$d(P) \triangleq \text{argmax}_d E_P E_P(Y|A = d(W), W)$$

is an optimal treatment rule under P. We will resolve the ambiguity in the definition of d when the argmax is not unique later in this section. We can identify $d(P)$ with a causally optimal rule under causal assumptions Pearl (2009). As the focus of this chapter is statistical in nature, all of the results will hold when estimating the parameter $\Psi(P_0)$ whether or not the causal assumptions needed for identifiability hold.

Consider the efficient influence curve of Ψ_d at P:

$$D(d, P)(O) = \frac{I(A = d(W))}{g(A|W)}(Y - \bar{Q}(A, W)) + \bar{Q}(d(W), W) - \Psi_d(P).$$

Let $B(P) \triangleq \{w : \bar{Q}_b(w) = 0\}$. We will refer to $B(P_0)$ as B_0. An exceptional law is defined as a distribution P for which $\Pr_P(W \in B(P)) > 0$ (Robins, 2004). We note that the ambiguity in the definition of $d(P)$ occurs precisely on the set $B(P)$. In particular, $d(P)$ must almost surely agree with some rule in the class

$$\left\{w \mapsto I(\bar{Q}_b(w) > 0)I(w \notin B(P)) + b(w)I(w \in B(P)) : b\right\}, \tag{2.1}$$

where $b : \mathcal{W} \to \{0, 1\}$ is some function. Consider now the following uniquely defined optimal rule:

$$d^*(P)(W) \triangleq I(\bar{Q}_b(W) > 0).$$

We will let $d_0^* = d^*(P_0)$. We have $\Psi(P) = \Psi_{d^*(P)}(P)$, but now $d^*(P)$ is uniquely defined for all W. More generally, $d^*(P)$ represents a uniquely defined optimal rule. Other formulations of the optimal rule can be obtained by changing the behavior of the rule B_0. Our goal is to construct root-n rate CIs for $\Psi(P_0)$ that maintain nominal coverage, even at exceptional laws. At non-exceptional laws we would like these CIs to belong to and be asymptotically efficient among the class of regular asymptotially linear (RAL) estimators.

2.3 Conditions for pathwise differentiability of the optimal value parameter

In this section we give a necessary and sufficient condition for the pathwise differentiability of the optimal value parameter Ψ. When it exists, the pathwise derivative in a nonparametric model can be written as an inner product between an almost surely unique mean zero, square integrable function known as the canonical gradient and a score function. The canonical gradient is a key object in nonparametric statistics. We remind the reader that an estimator $\hat{\Phi}$ is asymptotically linear for a parameter mapping Φ at P_0 with influence curve IC_0 if

$$\hat{\Phi}(P_n) - \Phi(P_0) = \frac{1}{n}\sum_{i=1}^{n} IC_0(O_i) + o_{P_0}(n^{-1/2}),$$

where $E_0[IC_0(O)] = 0$. The pathwise derivative is important because, when Φ is pathwise differentiable in a nonparametric model, any regular estimator $\hat{\Phi}$ is asymptotically linear with influence curve $IC_0(O_i)$ if and only if IC_0 is the canonical gradient (Bickel et al., 1993). We discuss negative results for non-pathwise differentiable parameters and formally define "regular estimator" later in this section.

The pathwise derivative of Ψ at P_0 can be defined as follows. Define paths $\{P_\epsilon : \epsilon \in \mathbb{R}\} \subset \mathcal{M}$ that go through P_0 at $\epsilon = 0$, i.e. $P_{\epsilon=0} = P_0$. In particular, these paths are given by

$$\begin{aligned} dQ_{W,\epsilon} &= (1 + \epsilon S_W(W))dQ_{W,0}, \\ &\quad \text{where } E_0[S_W(W)] = 0 \text{ and } \sup_w |S_W(w)| < \infty; \\ dQ_{Y,\epsilon}(Y|A,W) &= (1 + \epsilon S_Y(Y|A,W))dQ_{Y,0}(Y|A,W), \\ &\quad \text{where } E_0[S_Y|A,W] = 0 \text{ and } \sup_{w,a,y} |S_Y(y|a,w)| < \infty. \end{aligned} \tag{2.2}$$

Above $Q_{W,0}$ and $Q_{Y,0}$ are respectively the marginal distribution of W and the conditional distribution of Y given A, W under P_0. The parameter Ψ is not sensitive to fluctuations of $g_0(a|w) = \Pr_0(a|w)$, and thus we do not need to fluctuate this portion of the likelihood. The parameter Ψ is called pathwise differentiable at P_0 if

$$\frac{d}{d\epsilon}\Psi(P_\epsilon)\bigg|_{\epsilon=0} = \int D^*(P_0)(o)(S_W(w) + S_{Y|A,W}(y|a,w))dP_0(o)$$

for some P_0 mean zero, square integrable function $D^*(P_0)$ with $E_0[D^*(P_0)(O)|A,W]$ almost surely equal to $E_0[D^*(P_0)(O)|W]$. We refer the reader to Bickel et al. (1993) for a more general exposition of pathwise differentiability.

In van der Laan and Luedtke (2014b), we showed that Ψ is pathwise differentiable at P_0 with canonical gradient $D(d_0^*, P_0)$ if P_0 is a non-exceptional law, i.e. $\Pr_0(W \notin B_0) = 1$. Exceptional laws were shown to present problems for estimation of optimal rules indexed by a finite dimensional parameter by Robins (2004), and it was observed by Robins and Rotnitzky (2014) that these laws can also cause problems for unrestricted optimal rules. Here we show that mean outcome under the optimal rule is pathwise differentiable under a slightly more general condition than requiring a non-exceptional law, namely that

$$\Pr_0\left\{w \in \mathcal{W} : w \notin B_0 \text{ or } \max_{a\in\{0,1\}} \sigma_0^2(a,w) = 0\right\} = 1, \tag{2.3}$$

where $\sigma_0(a,w) \triangleq \sqrt{\mathrm{Var}_{P_0}(Y|A=a, W=w)}$. The upcoming theorem also gives the converse result, i.e. the mean outcome under the optimal rule is not pathwise differentiable if the above condition does not hold.

Theorem 1. *Assume* $\Pr_0(0 < g_0(1|W) < 1) = 1$, $\Pr_0(|Y| < M) = 1$ *for some* $M < \infty$, *and* $\mathrm{Var}_{P_0}(D(d_0^*, P_0)(O)) < \infty$. *The parameter* $\Psi(P_0)$ *is pathwise differentiable if and only if (2.3) holds. If* Ψ *is pathwise differentiable at* P_0, *then* Ψ *has canonical gradient* $D(d_0^*, P_0)$ *at* P_0.

In the proof of the theorem we construct fluctuations S_W and S_Y such that

$$\lim_{\epsilon\uparrow 0} \frac{\Psi(P_\epsilon) - \Psi(P_0)}{\epsilon} \neq \lim_{\epsilon\downarrow 0} \frac{\Psi(P_\epsilon) - \Psi(P_0)}{\epsilon} \tag{2.4}$$

when (2.3) does not hold. It then follows that $\Psi(P_0)$ is not pathwise differentiable. The left- and right-hand sides above are referred to as one-sided directional derivatives by Hirano and Porter (2012).

This condition for the mean outcome differs slightly from that implied for unrestricted rules in Robins and Rotnitzky (2014) in that we still have pathwise differentiability when the $\bar{Q}_{b,0}$ is zero in some strata but the conditional variance of the outcome given covariates and treatment is also zero in all of those strata. This makes sense, given that in this case the blip function could be estimated perfectly in such strata in any finite sample with treated and untreated individuals observed in those strata. Though we do not expect this difference to matter for most data generating distributions encountered in practice, there are cases where it may be relevant. For example, if no one in a certain stratum is susceptible to a disease regardless of treatment status, and researchers are unaware of this *a priori* so that simply excluding this stratum from the target population is not an option, then the treatment effect and conditional variance are both zero in this stratum.

In general, however, we expect that the mean outcome under the optimal rule will not be pathwise differentiable under exceptional laws encountered in practice. For this reason, we often refer to "exceptional laws" rather than "laws which do not satisfy (2.3)". We do this because the term "exceptional law" is well-established in the literature, and also because we believe that there is likely little distinction between "exceptional law" and "laws which do not satisfy (2.3)" for many problems of interest.

For the definitions of regularity and local unbiasedness we let P_ϵ be as in (2.2), with g_0 also fluctuated. That is, we let $dP_\epsilon = dQ_{Y,\epsilon} \times g_\epsilon \times dQ_{W,\epsilon}$, where $g_\epsilon(A|W) = (1 + \epsilon S_A(A|W))g_0(A|W)$ with $E_0[S_A(A|W)|W] = 0$ and $\sup_{a,w} |S_A(a|w)| < \infty$. The estimator $\hat{\Phi}$ of $\Phi(P_0)$ is called regular if the asymptotic distribution of $\sqrt{n}(\hat{\Phi}(P_n) - \Phi(P_0))$ is not sensitive to small fluctuations in P_0. That is, the limiting distribution of $\sqrt{n}(\hat{\Phi}(P_{n,\epsilon=1/\sqrt{n}}) - \Phi(P_{\epsilon=1/\sqrt{n}}))$ does not depend on S_W, S_A, or S_Y, where $P_{n,\epsilon=1/\sqrt{n}}$ is the empirical distribution $O_1, ..., O_n$ drawn i.i.d. from $P_{\epsilon=1/\sqrt{n}}$. The estimator $\hat{\Phi}$ is called locally unbiased if the limiting distribution of $\sqrt{n}(\hat{\Phi}(P_{n,\epsilon=1/\sqrt{n}}) - \Phi(P_{\epsilon=1/\sqrt{n}}))$ has mean zero for all fluctuations S_W, S_A, and S_Y, and is called asymptotically unbiased (at P_0) if the bias of $\hat{\Phi}(P_n)$ for the parameter $\Phi(P_0)$ is $o_{P_0}(n^{-1/2})$ at P_0.

The non-regularity of a statistical inference problem does not typically imply the nonexistence of asymptotically unbiased estimators (see Example 4 of Liu and Brown, 1993 and the discussion thereof in Chen, 2004), but rather the nonexistence of *locally* asymptotically unbiased estimators whenever (2.4) holds for some fluctuation (Hirano and Porter, 2012). It is thus not surprising that we are able to find an estimator that is asymptotically unbiased at a fixed (possibly exceptional) law under mild assumptions. Hirano and Porter also show that there does not exist a regular estimator of the optimal value at any law for which (2.4) holds for some fluctuation. That is, no regular estimators of $\Psi(P_0)$ exist at laws which satisfy the conditions of Theorem 1 but do not satisfy (2.3), i.e. one must accept the non-regularity of their estimator when the data is generated from such a law. Note that this does not rule out the development of locally consistent confidence bounds similar to those presented by Laber and Murphy (2011) and Laber et al. (2014b), though such approaches can be conservative when the estimation problem is non-regular.

In this chapter we present an estimator $\hat{\Psi}$ for which $\Gamma_n\sqrt{n}(\hat{\Psi}(P_n) - \Psi(P_0))$ converges in distribution to a standard normal distribution for a random standardization term Γ_n under reasonable conditions. Our estimator does not require any complicated tuning parameters, and thus allows one to easily develop root-n rate CIs for the optimal value. We show that our estimator is RAL and efficient at laws which satisfy (2.3) under conditions.

2.4 Inference at exceptional laws

The challenge

Before presenting our estimator, we discuss the challenge of estimating the optimal value at exceptional laws. Suppose d_n is an estimate of d_0^* and $\hat{\Psi}_{d_n}(P_n)$ is an estimate of $\Psi(P_0)$ relying on the full data set. In van der Laan and Luedtke (2014a) we presented a targeted minimum loss-based estimator (TMLE) $\hat{\Psi}_{d_n}(P_n)$ which satisfies

$$\hat{\Psi}_{d_n}(P_n) - \Psi(P_0) = (P_n - P_0)D(d_n, P_n^*) + \underbrace{\Psi_{d_n}(P_0) - \Psi(P_0)}_{o_{P_0}(n^{-1/2}) \text{ under conditions}} + o_{P_0}(n^{-1/2}),$$

where we use the notation $Pf = E_P[f(O)]$ for any distribution P and the second $o_{P_0}(n^{-1/2})$ term is a remainder from a first-order expansion of Ψ. The term $\Psi_{d_n}(P_0) - \Psi(P_0)$ being $o_{P_0}(n^{-1/2})$ relies on the optimal rule being estimated well in terms of value and will often prove to be a reasonable condition, even at exceptional laws (see Theorem 5 in Section 2.7). Here P_n^* is an estimate of the components of P_0 needed to estimate $D(d_n, P_0)$. To show asymptotic linearity, one might try to replace $D(d_n, P_n^*)$ with a term that does not rely on the sample:

$$(P_n - P_0)D(d_n, P_n^*) = (P_n - P_0)D(d_0^*, P_0) + \underbrace{(P_n - P_0)(D(d_n, P_n^*) - D(d_0^*, P_0))}_{\text{empirical process}}.$$

If $D(d_n, P_n^*)$ belongs to a Donsker class and converges to $D(d_0^*, P_0)$ in $L^2(P_0)$, then the empirical process term is $o_{P_0}(n^{-1/2})$ and $\sqrt{n}(\hat{\Psi}_{d_n}(P_n) - \Psi(P_0))$ converges in distribution to a normal random variable with mean zero and variance $\mathrm{Var}_{P_0}(D(d_0^*, P_0))$ (van der Vaart and Wellner, 1996). Note that $D(d_n, P_n^*)$ being consistent for $D(d_0^*, P_0)$ will typically rely on d_n being consistent for the fixed d_0^* in $L^2(P_0)$, which we emphasize is *not* implied by $\Psi_{d_n}(P_0) - \Psi(P_0) = o_{P_0}(n^{-1/2})$. Zhang et al. (2012b) make this assumption in the regularity conditions in their Web Appendix A when they consider an analogous empirical process term in deriving the standard error of an estimate of the optimal value in a restricted class. More specifically, Zhang et al. assume a non-exceptional law and consistent estimation of a fixed optimal rule. Van der Laan and Luedtke (2014a) also make such an assumption. If P_0 is an exceptional law, then we likely do not expect d_n to be consistent for any fixed (non-data dependent) function. Rather, we expect d_n to fluctuate randomly on the set B_0, even as the sample size grows to infinity. In this case the empirical process term considered above is not expected to behave as $o_{P_0}(n^{-1/2})$.

Accepting that our estimates of the optimal rule may not stabilize as sample size grows, we consider an estimation strategy that allows d_n to remain random even as $n \to \infty$.

A thought experiment

First we give an erroneous estimation strategy which contains the main idea of the approach but is not correct in its current form. A modification is given in the next section. For simplicity, we will assume that one knows $v_n \triangleq \mathrm{Var}_{P_0}(D(d_n, P_0))$ given an estimate d_n and, for simplicity, that v_n is almost surely bounded away from zero. Under reasonable conditions,

$$v_n^{-1/2}\left(\hat{\Psi}_{d_n}(P_n) - \Psi(P_0)\right) = (P_n - P_0)v_n^{-1/2}D(d_n, P_n^*) + o_{P_0}(n^{-1/2}).$$

The empirical process on the right is difficult to handle because d_n and v_n are random quantities that likely will not stabilize to a fixed limit at exceptional laws.

As a thought experiment, suppose that we could treat $\{v_n^{-1/2} D(d_n, P_n^*) : n\}$ as a deterministic sequence, where this sequence does not necessarily stabilize as sample size grows. In this case the Lindeberg-Feller central limit theorem (CLT) for triangular arrays (see, e.g., Athreya and Lahiri, 2006) would allow us to show that the leading term on the right-hand side converges to a standard normal random variable. This result relies on inverse weighting by $\sqrt{v_n}$ so the variance of the terms in the sequence stabilizes to one as sample size gets large.

Of course we cannot treat these random quantities as deterministic. In the next section we will use the general trick of inverse weighting by the standard deviation of the terms over which we are taking an empirical mean, but we will account for the dependence of the estimated rule d_n on the data by inducing a martingale structure that allows us to treat a sequence of estimates of the optimal rule as known (conditional on the past). We can then apply a martingale CLT for triangular arrays to obtain a limiting distribution for our estimator.

2.5 Estimation of and inference for the optimal value

In this section we present a modified one-step estimator $\hat{\Psi}$ of the optimal value. This estimator relies on estimates of the treatment mechanism g_0, the stratum-specific outcome $\bar{Q}_0$, and the optimal rule d_0^*. We first present our estimator, and then present an asymptotically valid two-sided CI for the optimal value under conditions. Next we give conditions under which our estimator is RAL and efficient, and finally we present a (potentially conservative) asymptotically valid one-sided CI which lower bounds the mean outcome under the unknown optimal treatment rule. The one-sided CI uses the same lower bound from the two-sided CI, but does not require a condition about the rate at which the value of the optimal rule converges to the optimal value, or even that the value of the estimated rule is consistent for the optimal value.

The estimators in this section can be extended to a martingale-based TMLE for $\Psi(P_0)$. Because the primary purpose of this chapter is to deal with inference at exceptional laws, we will only present an online one-step estimator and leave the presentation of such a TMLE to future work.

Estimator of the optimal value

In this section we present our estimator of the optimal value. Our procedure first estimates the needed features g_0, $\bar{Q}_0$, and d_0^* of the likelihood based on a small chunk of data, and then evaluates a one-step estimator with these nuisance function values on the next chunk of the data. It then estimates the features on the first two chunks of data, and evaluates the one-step estimator on the next chunk of data. This procedure iterates until we have a sequence of estimates of the optimal value. We then output a weighted average of these chunk-specific estimates as our final estimate of the optimal value. While the first chunk needs to be large enough to estimate the desired nuisance parameters, i.e. large enough to estimate the features, all subsequent chunks can be of arbitrary size (as small as a single observation).

We now formally describe our procedure. Define

$$\tilde{D}(d, \bar{Q}, g)(o) \triangleq \frac{I(a = d(w))}{g(a|w)}(Y - \bar{Q}(a, w)) + \bar{Q}(d(w), w).$$

Let $\{\ell_n\}$ be some sequence of nonnegative integers representing the smallest sample on which the optimal rule is learned. For each $j = 1, ..., n$, let $P_{n,j}$ represent the empirical distribution of the observations $(O_1, O_2, ..., O_j)$. Let $g_{n,j}$, $\bar{Q}_{n,j}$, and $d_{n,j}$ respectively represent estimates of the g_0, $\bar{Q}_0$, and d_0^* based on (some subset of) the observations $(O_1, ..., O_{j-1})$ for all $j > \ell_n$. We subscript each of these estimates by both n and j because the subsets on which these estimates are obtained may depend on sample size. We give an example of a situation where this would be desirable in Section 2.6.

Define

$$\tilde{\sigma}^2_{0,n,j} \triangleq \mathrm{Var}_{P_0}\left(\tilde{D}(d_{n,j}, \bar{Q}_{n,j}, g_{n,j})(O) \middle| O_1, ..., O_{j-1}\right).$$

Let $\tilde{\sigma}^2_{n,j}$ represent an estimate of $\tilde{\sigma}^2_{0,n,j}$ based on (some subset of) the observations O_1 through O_{j-1}. Note that we omit the dependence of $\tilde{\sigma}_{n,j}$ and $\tilde{\sigma}_{0,n,j}$ on $d_{n,j}$, $\bar{Q}_{n,j}$, and $g_{n,j}$ in the notation. Our results apply to any sequence of estimates $\tilde{\sigma}^2_{n,j}$ which satisfies Conditions C1) through C5), which are stated later in this section. Also define

$$\Gamma_n \triangleq \frac{1}{n - \ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}^{-1}_{n,j}.$$

Our estimate $\hat{\Psi}(P_n)$ of $\Psi(P_0)$ is given by

$$\hat{\Psi}(P_n) \triangleq \Gamma_n^{-1} \frac{1}{n - \ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}^{-1}_{n,j} \tilde{D}_{n,j}(O_j) = \frac{\sum_{j=\ell_n+1}^{n} \tilde{\sigma}^{-1}_{n,j} \tilde{D}_{n,j}(O_j)}{\sum_{j=\ell_n+1}^{n} \tilde{\sigma}^{-1}_{n,j}}, \tag{2.5}$$

where $\tilde{D}_{n,j} \triangleq \tilde{D}(d_{n,j}, \bar{Q}_{n,j}, g_{n,j})$. We note that the Γ_n^{-1} standardization is used to account for the term-wise inverse weighting so that $\hat{\Psi}(P_n)$ estimates $\Psi(P_0) = E_0[\tilde{D}(d_0^*, \bar{Q}_0, g_0)]$. The above looks a lot like a standard augmented inverse probability weighted (AIPW) estimator, but with d_0^* estimated on chunks of data increasing in size and with each term in the sum given weight proportional to an estimate of the conditional variance of that term. Our estimator constitutes a minor modification of the online one-step estimator presented in van der Laan and Lendle (2014). In particular, each term in the sum is inverse weighted by an estimate of the standard deviation of $\tilde{D}_{n,j}$. For ease of reference we will refer to the estimator above as an online one-step estimator.

This estimation scheme differs from sample split estimation, where features are estimated on half of the data and then a one-step estimator is evaluated on the remaining half of the data. While one can show that such estimators achieve valid coverage using Wald-type CIs, these CIs will generally be approximately $\sqrt{2}$ times larger than the CIs of our proposed procedure (see the next section) because the one-step estimator is only applied to half of the data. Alternatively, one could try averaging two such estimators, where the training and the one-step sample are swapped between the two estimators. Such a procedure will fail to yield valid Wald-type CIs due to the non-regularity of the inference problem: one cannot replace the optimal rule estimates with their limits because such limits will not generally exist, and thus the estimator averages over terms with a complicated dependence structure.

Two-sided confidence interval for the optimal value

Define the remainder terms

$$
\begin{aligned}
R_{1n} &\triangleq \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\tilde{\sigma}_{n,j}^{-1}E_0\Bigg[\left(1-\frac{g_0(d_{n,j}(W)|W)}{g_{n,j}(d_{n,j}(W)|W)}\right) \\
&\qquad\qquad \times\left(\bar{Q}_{n,j}(d_{n,j}(W),W)-\bar{Q}_0(d_{n,j}(W),W)\right)\Bigg], \\
R_{2n} &\triangleq \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\frac{\Psi_{d_{n,j}}(P_0)-\Psi(P_0)}{\tilde{\sigma}_{n,j}}.
\end{aligned}
$$

The upcoming theorem relies on the following assumptions:

C1) $n-\ell_n$ diverges to infinity as n diverges to infinity.

C2) Lindeberg-like condition: for all $\epsilon > 0$,

$$
\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}E_0\left[\left(\frac{\tilde{D}_{n,j}(O)}{\tilde{\sigma}_{n,j}}\right)^2 T_{n,j}(O)\,\middle|\, O_1,...,O_{j-1}\right]=o_{P_0}(1),
$$

where $T_{n,j}(O)\triangleq I\left(\frac{|\tilde{D}_{n,j}(O)|}{\tilde{\sigma}_{n,j}}>\epsilon\sqrt{n-\ell_n}\right)$.

C3) $\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\frac{\tilde{\sigma}_{0,n,j}^2}{\tilde{\sigma}_{n,j}^2}$ converges to 1 in probability.

C4) $R_{1n}=o_{P_0}(n^{-1/2})$.

C5) $R_{2n}=o_{P_0}(n^{-1/2})$.

The assumptions are discussed in Section 2.7. We note that all of our results also hold with R_{1n} and R_{2n} behaving as $o_{P_0}(1/\sqrt{n-\ell_n})$, though we do not expect this observation to be of use in practice as we recommend choosing ℓ_n so that $n-\ell_n$ increases at the same rate as n.

Theorem 2. *Under Conditions C1) through C5), we have that*

$$\Gamma_n \sqrt{n-\ell_n}\left(\hat{\Psi}(P_n) - \Psi(P_0)\right) \rightsquigarrow N(0,1),$$

where we use "$\rightsquigarrow$" to denote convergence in distribution as the sample size converges to infinity. It follows that an asymptotically valid $1-\alpha$ CI for $\Psi(P_0)$ is given by

$$\hat{\Psi}(P_n) \pm z_{1-\alpha/2} \frac{\Gamma_n^{-1}}{\sqrt{n-\ell_n}},$$

where $z_{1-\alpha/2}$ denotes the $1-\alpha/2$ quantile of a standard normal random variable.

We have shown that, under very general conditions, the above CI yields an asymptotically valid $1-\alpha$ CI for $\Psi(P_0)$. We refer the reader to Section 2.7 for a detailed discussion of the conditions of the theorem. We note that our estimator is asymptotically unbiased, i.e. has bias of the order $o_{P_0}(n^{-1/2})$, provided that $\Gamma_n = O_{P_0}(1)$ and $n-\ell_n$ grows at the same rate as n.

Interested readers can consult the proof of Theorem 2 in Section 2.11 for a better understanding of why we proposed the particular estimator given in Section 2.5.

Conditions for asymptotic efficiency

We will now show that, if P_0 is a non-exceptional law and $d_{n,j}$ has a fixed optimal rule limit d_0, then our online estimator is RAL for $\Psi(P_0)$. The upcoming corollary makes use of the following consistency conditions for some fixed rule d_0 which falls in the class of optimal rules given in (2.1):

$$\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} E_0\left[(d_{n,j}(W) - d_0(W))^2 \middle| O_1, ..., O_{j-1}\right] = o_{P_0}(1) \tag{2.6}$$

$$\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} E_0\left[\left(\bar{Q}_{n,j}(d_0(W),W) - \bar{Q}_0(d_0(W),W)\right)^2 \middle| O_1, ..., O_{j-1}\right] = o_{P_0}(1) \tag{2.7}$$

$$\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} E_0\left[(g_{n,j}(d_0(W)|W) - g_0(d_0(W)|W))^2 \middle| O_1, ..., O_{j-1}\right] = o_{P_0}(1). \tag{2.8}$$

It also makes use of the following conditions, which are, respectively, slightly stronger than Conditions C1) and C3):

C1') $\ell_n = o(n)$.

C3') $\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\left|\frac{\tilde{\sigma}_{0,n,j}^2}{\tilde{\sigma}_{n,j}^2} - 1\right| \to 0$ in probability.

Corollary 1. *Suppose that Conditions C1'), C2), C3'), C4), and C5) hold. Also suppose that* $\Pr_0(\delta < g_0(1|W) < 1-\delta) = 1$ *for some* $\delta > 0$*, the estimates* $g_{n,j}$ *are bounded away from zero with probability* 1*,* Y *is bounded, the estimates* $\bar{Q}_{n,j}$ *are uniformly bounded,* $\ell_n = o(n)$*, and that, for some fixed optimal rule* d_0*, (2.6), (2.7), and (2.8) hold. Finally, assume that* $\text{Var}_{P_0}(\tilde{D}(d_0, \bar{Q}_0, g_0)) > 0$ *and that, for some* $\delta_0 > 0$*, we have that*

$$\Pr_0\left(\inf_{j,n} \tilde{\sigma}^2_{n,j} > \delta_0\right) = 1,$$

where the infimum is over natural number pairs (j, n) *for which* $\ell_n < j \leq n$*. Then we have that*

$$\Gamma_n^{-1} \to \text{Var}_{P_0}(\tilde{D}(d_0, \bar{Q}_0, g_0)) \text{ in probability as } n \to \infty. \tag{2.9}$$

Additionally,

$$\hat{\Psi}(P_n) - \Psi(P_0) = \frac{1}{n}\sum_{i=1}^{n} D(d_0, P_0) + o_{P_0}(1/\sqrt{n}). \tag{2.10}$$

That is, $\hat{\Psi}(P_n)$ *is asymptotically linear with influence curve* $D(d_0, P_0)$*. Under the conditions of this corollary, it follows that* P_0 *satisfies (2.3) if and only if* $\hat{\Psi}(P_n)$ *is RAL and asymptotically efficient among all such RAL estimators.*

We note that (2.9) combined with C1') implies that the CI given in Theorem 2 asymptotically has the same width (up to an $o_{P_0}(n^{-1/2})$ term) as the CI which treats (2.10) and $D(d_0, P_0)$ as known and establishes a typical Wald-type CI about $\hat{\Psi}(P_n)$.

The empirical averages over j in (2.6), (2.7), and (2.8) can easily be dealt with using Lemma 1, presented in Section 2.7. Essentially we have required that $d_{n,j}$, $\bar{Q}_{n,j}$, and $g_{n,j}$ are consistent for d_0, $\bar{Q}_0$, and g_0 as n and j get large, where d_0 is some fixed optimal rule. One would expect such a fixed limiting rule d_0 to exist at a non-exceptional law for which the optimal rule is (almost surely) unique. If g_0 is known then we do not need $\bar{Q}_{n,j}$ to be consistent for $\bar{Q}_0$ to get asymptotic linearity, but rather that $\bar{Q}_{n,j}$ converges to some possibly misspecified fixed limit $\bar{Q}$.

Lower bound for the optimal value

It would likely be useful to have a conservative lower bound on the optimal value in practice. If policymakers were to implement an optimal individualized treatment rule whenever the overall benefit is greater than some fixed threshold, i.e. $\Psi(P_0) > v$ for some fixed v, then a one-sided CI for $\Psi(P_0)$ would help facilitate the decision to implement an individualized treatment strategy in the population.

The upcoming theorem shows that the lower bound from the $1-2\alpha$ CI yields a (potentially conservative) asymptotic $1-\alpha$ CI for the optimal value. If d_0^* is estimated well in the sense of Condition C5), then the asymptotic coverage is exact. Define

$$LB_n(\alpha) \triangleq \hat{\Psi}(P_n) - z_{1-\alpha}\frac{\Gamma_n^{-1}}{\sqrt{n-\ell_n}}.$$

Theorem 3. *Under Conditions C1) through C4), we have that*

$$\liminf_{n\to\infty} \Pr_0\left(\Psi(P_0) > LB_n(\alpha)\right) \geq 1-\alpha.$$

If Condition C5) also holds, then

$$\lim_{n\to\infty} \Pr_0\left(\Psi(P_0) > LB_n(\alpha)\right) = 1-\alpha.$$

The above condition should not be surprising, as we base our CI for $\Psi(P_0)$ on a weighted combination of estimates of $\Psi_{d_{n,j}}(P_0)$ for $j < n$. Because $\Psi(P_0) \geq \Psi_{d_{n,j}}(P_0)$ for all such j, we would expect that the lower bound of the $1-\alpha$ CI given in the previous section provides a valid $1-\alpha/2$ one-sided CI for $\Psi(P_0)$. Indeed this is precisely what we see in the proof of the above theorem.

2.6 Computationally efficient estimation schemes

Computing $\hat{\Psi}(P_n)$ may initially seem computationally demanding. In this section we discuss two estimation schemes which yield computationally simple routines.

Computing the features on large chunks of the data

One can compute the estimates of $\bar{Q}_0$, g_0, and d_0^* far fewer than $n - \ell_n$ times. For each j, the estimates $\bar{Q}_{n,j}$, $g_{n,j}$, and $d_{n,j}$ may rely on any subset of the observations $O_1, ..., O_{j-1}$. Thus one can compute these estimators on S increasing subsets of the data, where the first subset consists of observations $O_1, ..., O_{\ell_n}$ and each of the $S-1$ remaining samples adds a $1/S$ proportion of the remaining $n-\ell_n$ observations. Note that this scheme makes use of the fact that, for fixed j, the feature estimates, indexed by n and j, e.g. $d_{n,j}$, may rely on different subsets of observations $O_1, ..., O_{j-1}$ for different sample sizes n.

Online learning of the optimal value

Our estimator was inspired by online estimators which can operate on large data sets that will not fit into memory. These estimators use online prediction and regression algorithms which update the intial fit based on previously observed estimates using new observations as they are read into memory. Online estimators of pathwise differentiable parameters were introduced in van der Laan and Lendle (2014). Such estimation procedures often require estimates of features of the likelihood, which can be obtained using modern online regression and classification approaches (see, e.g., Zhang, 2004; Langford et al., 2009; Luts et al., 2014). Our estimator constitutes a slight modification of the one-step online estimator presented by van der Laan and Lendle (2014), and thus all discussion of computational efficiency given in that chapter applies to our case.

For our estimator, one could use online estimators of $\bar{Q}_0$, g_0, and d_0^*, and then update these estimators as the index j in the sum in (2.5) increases. Calculating the standard error estimate $\tilde{\sigma}_{n,j}$ will typically require access to an increasing subset of the past observations, i.e. as sample size grows one may need to hold a growing number of observations in memory. If one uses a sample standard deviation to estimate $\tilde{\sigma}_{0,n,j}$ based on subset of observations $O_1, ..., O_{j-1}$, the results we present in Section 2.7 will indicate that one really only needs that the number of points on which $\tilde{\sigma}_{0,n,j}$ is estimated grows with j rather than at the same rate as j. This suggest that, if computation time or system memory is a concern for calculating $\tilde{\sigma}_{n,j}$, then one could calculate $\tilde{\sigma}_{n,j}$ based on some $o(j)$ subset of observations $O_1, ..., O_{j-1}$.

2.7 Discussion of the conditions of Theorem 2

For ease of notation we will assume that, for all $j > \ell_n$, we do not modify our feature estimates based on the first $j-1$ data points as the sample size grows. That is, for all sample sizes m, n and all $j \leq \min\{m, n\}$, $d_{n,j} = d_{m,j}$, $\bar{Q}_{n,j} = \bar{Q}_{m,j}$, $g_{n,j} = g_{m,j}$, and $\tilde{\sigma}_{n,j} = \tilde{\sigma}_{m,j}$. One can easily extend all of the discussion in this section to a more general case where, e.g., $d_{n,j} \neq d_{m,j}$ for $n \neq m$. This may be useful if the optimal rule is estimated in chunks of increasing size as was discussed in Section 2.6. To make these object's lack of dependence on n clear, in this section we will denote $d_{n,j}$, $\bar{Q}_{n,j}$, $g_{n,j}$, $\tilde{\sigma}_{n,j}$, and $\tilde{\sigma}_{0,n,j}$ as d_j, $\bar{Q}_j$, g_j, $\tilde{\sigma}_j$ and $\tilde{\sigma}_{0,j}$. This will also help make it clear when o_{P_0} notation refers to behavior as j, rather than n, goes to infinity.

For our discussion we assume there exists a (possibly unknown) $\delta_0 > 0$ such that

$$\Pr_0 \left(\inf_{j > \ell_n} \tilde{\sigma}_{0,j}^2 > \delta_0 \right) = 1, \tag{2.11}$$

where the probability statement is over the i.i.d. draws $O_1, O_2, ...$. The above condition is not necessary, but will make our discussion of the conditions more straightforward.

Discussion of Condition C1)

We cannot apply the martingale CLT in the proof of Theorem 2 if $n - \ell_n$ does not grow with sample size. Essentially this condition requires that a non-negligible proportion of the data is used to actually estimate the mean outcome under the optimal rule. One option is to have $n - \ell_n$ grow at the same rate as n grows, which holds if e.g. $\ell_n = pn$ for some fixed proportion p of the data. This allows our CIs to shrink at a root-n rate. One might prefer to have $\ell_n = o(n)$ so that $\frac{n-\ell_n}{n}$ converges to 1 as sample size grows. In this case we can show that our estimator is asymptotically linear and efficient at non-exceptional laws under conditions, as we did in Corollary 1.

Discussion of Condition C2)

This is a standard condition that yields a martingale CLT for triangular arrays (Gaenssler et al., 1978). The condition ensures that the variables which are being averaged have sufficiently thin tails. While it is worth stating the condition in general, it is easy to verify that the condition is implied by the following three more straightforward conditions:

- (2.11) holds.
- Y is a bounded random variable.
- There exists some $\delta > 0$ such that $\Pr_0(\delta < g_j(1|W) < 1-\delta) = 1$ with probability 1 for all j.

Indeed, under the latter two conditions $|\tilde{D}_{n,j}(O)| < C$ is almost surely bounded for some $C > 0$, and thus (2.11) yields that $|\tilde{D}_{n,j}(O)\tilde{\sigma}_{n,j}^{-1}| < C\delta_0^{-1} < \infty$ with probability 1. For all $\epsilon > 0$, $\epsilon\sqrt{n-\ell_n} > C\delta_0^{-1}$ for all n large enough under Condition C1). Thus $T_{n,j}$ from Condition C2) is equal to zero with probability 1 for all n large enough.

Discussion of Condition C3)

This is a rather weak condition given that $\tilde{\sigma}_{0,j}$ still treats d_j as random. Thus this condition does not require that d_j stabilizes as j gets large. Suppose that

$$\tilde{\sigma}_j^2 - \tilde{\sigma}_{0,j}^2 = o_{P_0}(1) \tag{2.12}$$

By (2.11) and the continuous mapping theorem, it follows that

$$\frac{\tilde{\sigma}_{0,j}^2}{\tilde{\sigma}_j^2} - 1 = o_{P_0}(1). \tag{2.13}$$

The following general lemma will be useful in establishing Conditions C3), C4), and C5).

Lemma 1. *Suppose that R_j is some sequence of (finite) real-valued random variables such that $R_j = o_{P_0}(j^{-\beta})$ for some $\beta \in [0,1)$, where we assume that each R_j is measurable with respect to the sigma-algebra generated by $(O_1, ..., O_j)$. Then,*

$$\frac{1}{n}\sum_{j=1}^{n} R_j = o_{P_0}(n^{-\beta}).$$

Applying the above lemma with $\beta = 0$ to (2.13) shows that Condition C3) holds provided that (2.11) and (2.12) hold. We will use the above lemma with $\beta = 1/2$ when discussing Conditions C4) and C5).

It remains to show that we can construct a sequence of estimators such that (2.12) holds. Suppose we estimate $\tilde{\sigma}^2_{0,j}$ with

$$\tilde{\sigma}^2_j \triangleq \max\left\{\delta_j, \frac{1}{j-1}\sum_{i=1}^{j-1}\tilde{D}^2_j(O_i) - \left(\frac{1}{j-1}\sum_{i=1}^{j-1}\tilde{D}_j(O_i)\right)^2\right\}, \tag{2.14}$$

where $\{\delta_j\}$ is a sequence that may rely on j and each $\tilde{D}_{n,j} = \tilde{D}_j$ for all $n \geq j$. We use δ_j to ensure that $\tilde{\sigma}^{-2}_j$ is well-defined (and finite) for all j. If a lower bound δ_0 on $\tilde{\sigma}^2_{0,j}$ is known then one can take $\delta_j = \delta_0$ for all j. Otherwise one can let $\{\delta_j\}$ be some sequence such that $\delta_j \downarrow 0$ as $j \to \infty$.

Note that $\tilde{\sigma}^2_j$ is an empirical process because it involves sums over observations O_1 through O_{j-1}, and functions $\tilde{D}_j$ which were estimated on those same observations. The following theorem gives sufficient conditions for (2.12), and thus Condition C3), to hold.

Theorem 4. *Suppose (2.11) holds and that $\left\{\tilde{D}(d, \bar{Q}, g) : d, \bar{Q}, g\right\}$ is a P_0 Glivenko-Cantelli (GC) class with an integrable envelope function, where d, $\bar{Q}$, and g are allowed to vary over the range of the estimators of d^*_0, $\bar{Q}_0$, and g_0. Let $\tilde{\sigma}^2_j$ be defined as in (2.14). Then we have that $\tilde{\sigma}^2_j - \tilde{\sigma}^2_{0,j} = o_{P_0}(1)$. It follows that (2.13) and Condition C3) are satisfied.*

We thus only make the very mild assumption that our estimators of d^*_0, $\bar{Q}_0$, and g_0 belong to GC classes. Note that this assumption is much milder than the typical Donsker condition needed when attempting to establish the asymptotic normality of a (non-online) one-step estimator. An easy sufficient condition for a class to have a finite envelope function is that it is uniformly bounded, which occurs if the conditions discussed in Section 2.7 hold.

Discussion of Condition C4)

This condition is a weighted version of the typical double robust remainder appearing in the analysis of the AIPW estimator. Suppose that

$$E_0\left[\left(1 - \frac{g_0(d_j(W)|W)}{g_j(d_j(W)|W)}\right)\left(\bar{Q}_j(d_j(W), W) - \bar{Q}_0(d_j(W), W)\right)\right] = o_{P_0}(j^{-1/2}). \tag{2.15}$$

If g_0 is known (as in an RCT without missingness) and one takes each $g_j = g_0$ then the above ceases to be a condition as the left-hand side is always zero. We note that the only condition on $\bar{Q}_j$ appears in Condition C4), so that if $R_{1n} = 0$ as in an RCT without missingness then we do not require that $\bar{Q}_j$ stabilizes as j grows. A typical AIPW estimator require the estimate of $\bar{Q}_0$ to stabilize as sample size grows to get valid inference, but here we have avoided this condition in the case where g_0 is known by using the martingale structure and inverse weighting by the standard error of each term in the definition of $\hat{\Psi}(P_n)$.

More generally, Lemma 1 shows that Condition C4) holds if (2.13) and (2.15) hold and $\Pr_0(0 < g_j(1|W) < 1) = 1$ with probability 1 for all j. One can apply the Cauchy-Schwarz inequality and take the maximum over treatment assignments to see that (2.15) holds if

$$\max\left\{\frac{\|g_j(a|W) - g_0(a|W)\|_{2,P_0}\left\|\bar{Q}_j(a,W) - \bar{Q}_0(a,W)\right\|_{2,P_0}}{g_j(a|W)} : a = 0,1\right\}$$

is $o_{P_0}(j^{-1/2})$. If g_0 is not known, the above shows that then (2.15) holds if g_0 and $\bar{Q}_0$ are estimated well.

Discussion of Condition C5)

This condition requires that we can estimate d_0^* well as sample size gets large. We now give a theorem which will help us to establish Condition C5) under reasonable conditions. The theorem assumes the following margin assumption: for some $\alpha > 0$,

$$\Pr_0(0 < |\bar{Q}_{b,0}(W)| \leq t) \lesssim t^{\alpha} \quad \forall t > 0, \tag{2.16}$$

where "$\lesssim$" denotes less than or equal to up to a nonnegative constant. This assumption is a direct restatement of Assumption (MA) from Audibert and Tsybakov (2007) and was considered earlier by Tsybakov (2004). Note that this theorem is similar in spirit to Lemma 1 in van der Laan and Luedtke (2014a), but relies on weaker, and we believe more interpretable, assumptions.

Theorem 5. *Suppose (2.16) holds for some $\alpha > 0$ and that we have an estimate $\bar{Q}_{b,n}$ of $\bar{Q}_{b,0}$ based on a sample of size n. If $\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0} = o_{P_0}(1)$, then*

$$|\Psi_{d_n}(P_0) - \Psi_{d_0^*}(P_0)| \lesssim \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0}^{2(1+\alpha)/(2+\alpha)},$$

where d_n is the function $w \mapsto I(\bar{Q}_{b,n}(w) > 0)$. If $\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0} = o_{P_0}(1)$, then

$$\begin{aligned}
&|\Psi_{d_n}(P_0) - \Psi_{d_0^*}(P_0)| \\
&\qquad \leq \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0} \Pr_0\left(0 < |\bar{Q}_{b,0}(W)| \leq \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0}\right) \\
&\qquad \lesssim \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0}^{1+\alpha}.
\end{aligned}$$

The above theorem thus shows that $\Psi_{d_j}(P_0) - \Psi_{d_0^*}(P_0) = o_{P_0}(j^{-1/2})$ the distribution of $|\bar{Q}_{b,0}(W)|$ and our estimates of $\bar{Q}_{b,0}$ satisfy reasonable conditions. If additionally $\tilde{\sigma}_{0,j}$ is estimated well in the sense of (2.13), then an application of Lemma 1 shows that Condition C5) is satisfied.

The first part of the proof of Theorem 5 is essentially a restatement of Lemma 5.2 in Audibert and Tsybakov (2007). Figure 2.1 shows various densities which satisfy (2.16) at

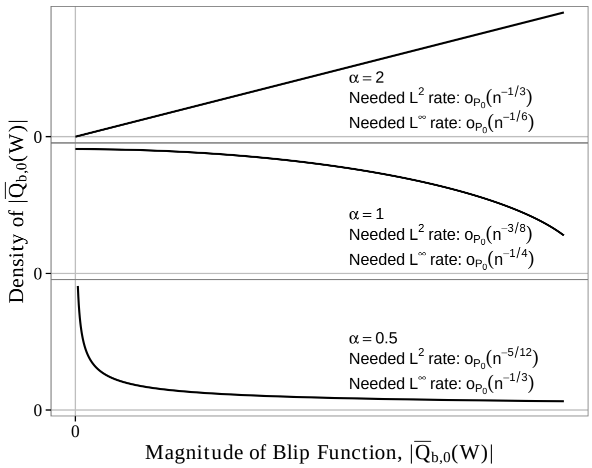

Figure 2.1: Examples of three densities of $|\bar{Q}_{b,0}(W)|$ whose corresponding cumulative distribution functions satisfy (2.16). If the rate of convergence of $\bar{Q}_{b,n} - \bar{Q}_{b,0}$ to zero in $L^2(P_0)$ or $L^\infty(P_0)$ attains the rates indicated above indicated above, then Condition C5) will be satisfied for the plug-in optimal rule estimate considered in Theorem 5.

different values of α, and also the slowest rate of convergence for the blip function estimates for which Theorem 5 implies Condition C5). As illustrated in the figure, $\alpha > 1$ implies that $p_{b,0}(t) \to 0$ as $t \to 0$. Given that we are interested in laws where $\mathrm{Pr}_0(\bar{Q}_{b,0}(W) = 0) > 0$, it is unclear how likely we are to have that $\alpha > 1$ when W contains only continuous covariates. One might, however, believe that the density is bounded near zero so that (2.16) is satisfied at $\alpha = 1$.

If $\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0} = o_{P_0}(1)$ then the above theorem indicates an arbitrarily fast rate for $\Psi_{d_n}(P_0) - \Psi_{d_0^*}(P_0)$ when there is a margin around zero, i.e. $\mathrm{Pr}_0(0 < |\bar{Q}_{b,0}(W)| \leq t) = 0$ for some $t > 0$. In fact, $\Psi_{d_n}(P_0) - \Psi_{d_0^*}(P_0) = 0$ with probability approaching 1 in this case. Such a margin will exist when W is discrete.

2.8 Simulation methods

We ran four simulations. Simulation D-E is a point treatment case, where the treatment may rely on a single categorical covariate W. Simulations C-NE and C-E are two different point treatment simulations where the treatment may rely on a single continuous covariate W. Simulation C-NE uses a non-exceptional law, while simulation C-E uses an exceptional law.

Each simulation setting was run over 2000 Monte Carlo draws to evaluate the performance of our new martingale-based method and a classical (and for exceptional laws incorrect) one-step estimator with Wald-type CIs. Table 2.1 shows the combinations of sample size (n) and

Simulation	(n, ℓ_n)
D-E	$(1000, 100)$, $(4000, 100)$
C-NE, C-E	$(250, 25)$, $(1000, 25)$, $(4000, 100)$

Table 2.1: Primary combinations of sample size (n) and initial chunk size (ℓ_n) considered in each simulation. Different choices of ℓ_n were considered for C-NE and C-E to explore the sensitivity of the estimator to the choice of ℓ_n.

initial chunk size (ℓ_n) considered for each estimator. All simulations were run in R (R Core Team, 2014).

Simulation D-E: discrete W

Data

This simulation uses a discrete baseline covariate W with four levels, a dichotomous treatment A, and a binary outcome Y. The data is generated by drawing i.i.d. samples from the distribution with $W \sim \text{Uniform}\{0, 1, 2, 3\}$, $A|W \sim \text{Binomial}(0.5 + 0.1W)$, and $Y|A, W \sim \text{Binomial}\left(0.4 + 0.2I(A = 1, W = 0)\right)$. This is an exceptional law because $\bar{Q}_{b,0}(w) = 0$ for $w \neq 0$. The optimal value is 0.45.

Estimation methods

For each $j = \ell_n+1, ..., n$, we used the nonparametric maximum likelihood estimator generated by the first $j-1$ samples to estimate P_0 and the corresponding plug-in estimators to estimate all of the needed features of the likelihood, including the optimal rule. We used the sample standard deviation of $\tilde{D}_{n,j}(O_1), ..., \tilde{D}_{n,j}(O_{j-1})$ to estimate $\tilde{\sigma}_{0,j}$.

Simulations C-NE and C-E: continuous univariate W

Data

This simulation uses a single continuous baseline covariate $W \sim \text{Uniform}(-1, 1)$ and dichotomous treatment A with conditional distribution $A|W \sim \text{Binomial}(0.5 + 0.1W)$. We consider two distributions for the binary outcome Y. The first distribution (C-NE) is a non-exceptional law with $Y|A, W$ drawn from a binomial distribution with probability of success $\bar{Q}_0^{\text{n-e}}(A, W)$, where

$$\bar{Q}_0^{\text{n-e}}(A, W) - \frac{3}{10} \triangleq \begin{cases} -W^3 + W^2 - \frac{1}{3}W + \frac{1}{27}, & \text{if } A = 1 \text{ and } W \geq 0 \\ \frac{3}{4}W^3 + W^2 - \frac{1}{3}W + \frac{1}{27}, & \text{if } A = 1 \text{ and } W < 0 \\ 0, & \text{if } A = 0. \end{cases}$$

The optimal value of approximately 0.388 was estimated using 10^8 Monte Carlo draws. The second distribution (C-E) is an exceptional law with $Y|A,W$ drawn from a binomial distribution with probability of success $\bar{Q}_0^{\mathrm{e}}(A,W)$, where for $\tilde{W} \triangleq W + 5/6$ we define

$$\bar{Q}_0^{\mathrm{e}}(A,W) - \frac{3}{10} \triangleq \begin{cases} -\tilde{W}^3 + \tilde{W}^2 - \frac{1}{3}\tilde{W} + \frac{1}{27}, & \text{if } A = 1 \text{ and } W < -1/2 \\ -W^3 + W^2 - \frac{1}{3}W + \frac{1}{27}, & \text{if } A = 1 \text{ and } W > 1/3 \\ 0, & \text{otherwise.} \end{cases}$$

The above distribution is an exceptional law because $\bar{Q}_0^{\mathrm{e}}(1,w) - \bar{Q}_0^{\mathrm{e}}(0,w) = 0$ whenever $w \in \left[-\frac{1}{2}, \frac{1}{3}\right]$. The optimal value of approximately 0.308 was estimated using 10^8 Monte Carlo draws.

Estimation methods

To show the flexibility of our estimation procedure with respect to estimators of the optimal rule, we estimated the blip functions using a Nadaraya-Watson estimator, where we behave as though g_0 is unknown when computing the kernel estimate. For the next simulation setting we use the ensemble learner from Luedtke and van der Laan (2014a) that we suggest using in practice. Here we estimated

$$\bar{Q}_{b,n}^h(w) \triangleq \frac{\sum_{i=1}^n y_i a_i K\left(\frac{w-w_i}{h}\right)}{\sum_{i=1}^n a_i K\left(\frac{w-w_i}{h}\right)} - \frac{\sum_{i=1}^n y_i (1-a_i) K\left(\frac{w-w_i}{h}\right)}{\sum_{i=1}^n (1-a_i) K\left(\frac{w-w_i}{h}\right)},$$

where $K(u) \triangleq \frac{3}{4}(1-u^2)I(|u| \le 1)$ is the Epanechnikov kernel and h is the bandwidth. Computing $\bar{Q}_{b,n}^h$ for a given bandwidth is the only point in our simulations where we do not treat g_0 as known. For a candidate blip function estimate $\bar{Q}_b$, define the loss

$$L_{\bar{Q}_0,g_0}(\bar{Q}_b)(o) \triangleq \left(\left[\frac{2a-1}{g_0(a|w)}(y - \bar{Q}_0(a,w)) + \bar{Q}_0(1,w) - \bar{Q}_0(0,w)\right] - \bar{Q}_b(w)\right)^2.$$

To save computation time we behave as though $\bar{Q}_0$ and g_0 are known when using the above loss. We selected the bandwidth H_n using 10-fold cross-validation with the above loss function to select from the candidates $h = (0.01, 0.02, ..., 0.20)$. We also behave as though $\bar{Q}_0$ and g_0 are known when estimating each $\tilde{D}_{n,j}$, so that the function $\tilde{D}_{n,j}$ only depends on $O_1, ..., O_{j-1}$ through the estimate of the optimal rule. This is mostly for convenience, as it saves on computation time and our estimate of the optimal rule d_0^* will still not stabilize, i.e. our optimal value estimators will still encounter the irregularity at exceptional laws. Note that g_0 is known in an RCT, and subtracting and adding $\bar{Q}_0$ in the definition of the loss function will only serve to stabilize the variance of our cross-validated risk estimate. In practice one could substitute an estimate of $\bar{Q}_0$ and expect similar results. We update our estimates $d_{n,j}$ and $\tilde{\sigma}_{0,n,j}$ using the method discussed in Section 2.6 with $S = \frac{n-\ell_n}{\ell_n}$.

To explore the sensitivity to the choice of ℓ_n we also considered (n, ℓ_n) pairs $(1000, 100)$ and $(4000, 400)$, where these pairs are only considered where explicitly noted. To explore the sensitivity of our estimators to permutations of our data, we ran our estimator twice on each Monte Carlo draw, with the indices of the observations permuted so that the online estimator sees the data in a different order.

2.9 Simulation results

Online one-step compared to classical one-step

Figure 2.2 shows the coverage attained by the online and classical (non-online) one-step estimates of the optimal value. The two-sided CIs resulting from the online estimator (nearly) attains nominal coverage for all simulations considered, whereas the non-online estimator only (nearly) attains nominal coverage for the non-exceptional law in C-NE. The one-sided CIs from the online one-step estimator attain proper coverage for all simulation settings. The one-sided CIs from the non-online one-step estimates do not (nearly) achieve nominal coverage in any of the simulations considered because the rule is estimated on the same data as the optimal value. Thus we expect to need a large sample size for the positive bias of the non-online one-step to be negligible. In van der Laan and Luedtke (2014a) we avoided this finite sample positive bias at non-exceptional laws by using a cross-validated TMLE for the optimal value.

Figure 2.3 displays the squared bias and mean CI length across the 2000 Monte Carlo draws. The online estimator consistently has lower squared bias across all of our simulations. The online estimator was negatively biased in all of our simulations, whereas the non-online estimator was positively biased in all of our simulations. This is not surprising: Theorem 3 already implies that the online estimator will generally be negatively biased in finite samples, whereas the non-online estimator will generally be positively biased as we have discussed.

2.10 Discussion

We have accomplished two tasks in this chapter. The first was to establish conditions under which we would expect that regular root-n rate inference is possible for the mean outcome under the optimal rule. In particular, we completely characterize the pathwise differentiability of the optimal value parameter. This characterization on the whole agrees with that implied by Robins and Rotnitzky (2014), but differs in a minor fringe case where the conditional variance of the outcome given covariates and treatment is zero. This fringe case may be relevant if everyone in a stratum of baseline covariates is immune to a disease (regardless of treatment status) but are still included in the study because experts are unaware of this immunity *a priori*. In general, however, the two characterizations agree.

The remainder of our work shows that one can obtain an asymptotically unbiased estimate of and a CI for the optimal value under reasonable conditions. This estimator uses a slight modification of the online one-step estimator presented by van der Laan and Lendle (2014). Under reasonable conditions, this estimator will be asymptotically efficient among all RAL

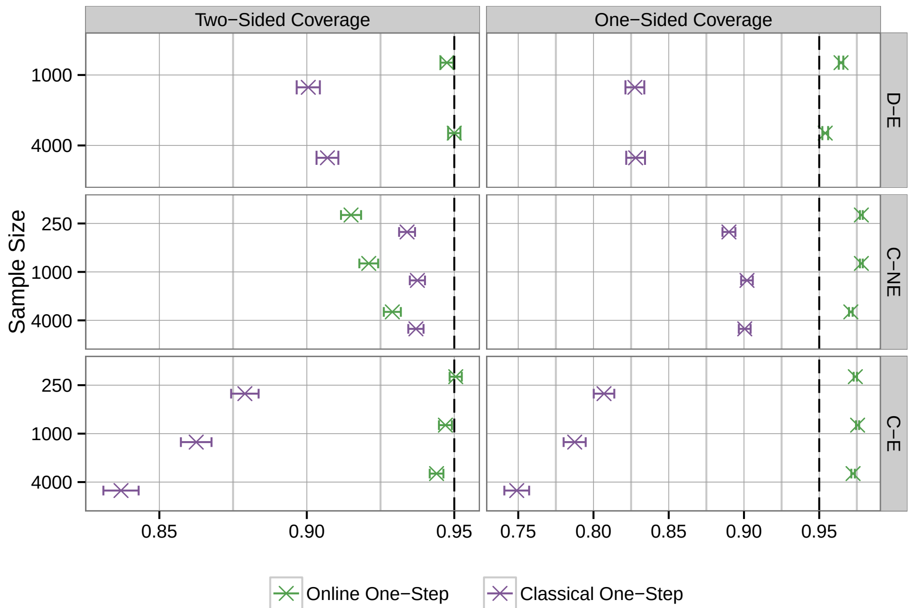

Figure 2.2: Coverage of 95% two-sided and one-sided (lower) CIs. The online one-step estimator achieves near-nominal coverage for all of the two-sided CIs and attains better than nominal coverage for the one-sided CI. The classical one-step estimator only achieves near-nominal coverage for C-NE. Error bars indicate 95% CIs to account for Monte Carlo uncertainty.

estimators of the optimal value at non-exceptional laws in the nonparametric model where the class of candidate treatment regimes is unrestricted. The main condition for the validity of our CI is that the value of one's estimate of the optimal rule converges to the optimal value at a faster than root-n rate, which we show is often a reasonable assumption. The lower bound in our CI is valid even if this condition does not hold.

We confirmed the validity of our approach using simulations. Our two-sided CIs attained near-nominal coverage for all simulation settings considered, while our lower CIs attained better than nominal coverage (they were conservative) for all simulation settings considered. Our CIs were of a comparable length to those attained by the non-online one-step estimator. The non-online one-step estimator only attained near-nominal coverage for the simulation which used a non-exceptional data generating distribution, as would be predicted by theory.

There is still more work to be done in estimating CIs for the optimal rule. While we have shown that the lower bound from our CI maintains nominal coverage under mild conditions, the upper bound requires the additional assumption that the optimal rule is estimated at a sufficiently fast rate. We observed in our simulations that the non-online estimate of the optimal value had positive bias for all settings. This is to be expected if the optimal rule is chosen to maximize the estimated value, and can easily be explained analytically under mild

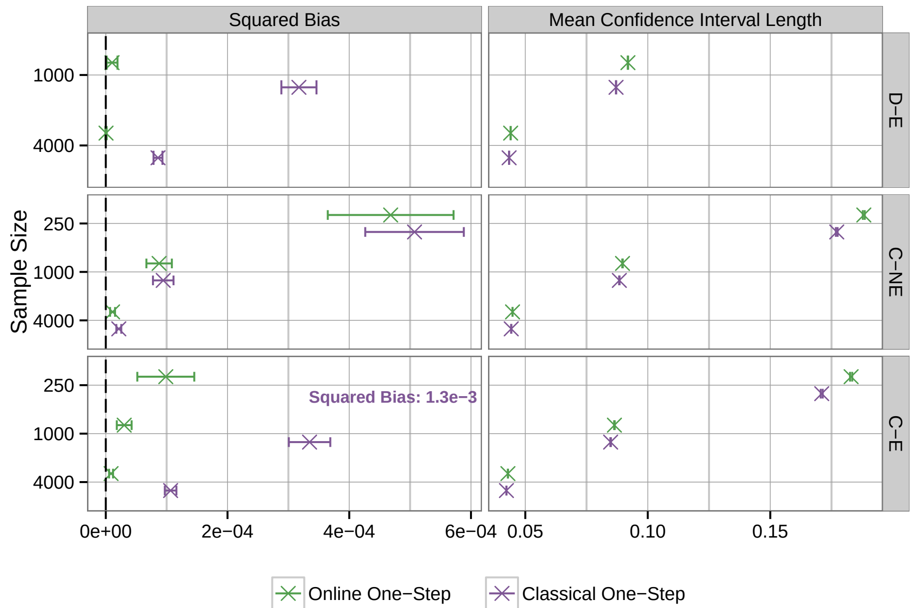

Figure 2.3: Squared bias and 95% two-sided CI lengths for the online and classical one-step estimators, where the mean is taken across 2000 Monte-Carlo draws. The online estimator has lower squared bias than the non-online estimator, while its mean CI length is only slightly longer. Error bars indicate 95% CIs to account for Monte Carlo uncertainty.

assumptions. It may be worth replacing the upper bound UB_n in our CI by something like $\max\{UB_n, \psi_n(d_n)\}$, where $\psi_n(d_n)$ is a non-online one-step estimate or TMLE of the optimal value. One might expect that the upper bound $\psi_n(d_n)$ will dominate the maximum precisely when the optimal rule is estimated poorly.

Our estimation strategy is not limited to unrestricted classes of optimal rules. One could replace our unrestricted class with, e.g., a parametric working model for the blip function and expect similar results. This is because the pathwise derivative of $P \mapsto E_{P_0}[Y_{d(P)}]$, which treats the P_0 in the expectation subscript as known, will typically be zero when $d(P)$ is an optimal rule in some class and does not fall on the boundary of that class (with respect to some metric). Such a result does not rely on $d(P)$ being a unique optimal rule. When the pathwise derivative of $P \mapsto E_{P_0}[Y_{d(P)}]$ is zero, one can often prove something like Theorem 5, which shows that the value of the estimated rule converges to the optimal value at a faster than root-n rate under conditions.

Here we considered the problem of developing a confidence interval for the value of an unknown optimal treatment rule in a non-parametric model. Under reasonable conditions, our proposed optimal value estimator provides an interpretable and statistically valid approach to gauging the effect of implementing the optimal individualized treatment regime in the population.

2.11 Proofs

Proofs of results from Section 2.3

Proof of Theorem 1. Let $d'(P)$ represent the function $w \mapsto I(\bar{Q}_b(P)(w) > 0)$. For any P, let $\Psi(P) \triangleq E_P E_P[Y|A = d(P)(W), W]$. Note that

$$\Psi(P) - E_P E_P[Y|A = 0, W] = E_P\left[d^*(P)(W)\bar{Q}_b(P)(W)\right] = E_P\left[d'(P)(W)\bar{Q}_b(P)(W)\right],$$

where we used the fact that $d^*(P)(w) = d'(P)(w)$ on the set where $\bar{Q}_b(P)(w) \neq 0$. Let the fluctuation submodel $\{P_\epsilon : \epsilon\}$ through P_0 be as defined in Section 2.3 of the main text, where we note that $P_0 = P_{\epsilon=0}$. Telescoping shows that, for fixed ϵ,

$$\Psi(P_\epsilon) - \Psi(P_0) = E_{P_\epsilon}\left[\left(I(\bar{Q}_{b,\epsilon} > 0) - I(\bar{Q}_{b,0} > 0)\right)\bar{Q}_{b,\epsilon}\right] + \Psi_{d_0'}(P_\epsilon) - \Psi_{d_0'}(P_0). \tag{2.17}$$

It is well known that $\Psi_d(P) \triangleq E_P E_P[Y|A = d(W), W]$ is pathwise differentiable for fixed d. Thus dividing the second line above by ϵ and taking the limit as $\epsilon \to 0$ yields the pathwise derivative that treats the rule d_0' as known. For a given S_Y, the fluctuated $\bar{Q}_{b,0}$ at $w \in \mathcal{W}$ is given by

$$\begin{aligned}\bar{Q}_{b,\epsilon}(w) &\triangleq \int y\left(dQ_{Y,\epsilon}(y|A = 1, W = w) - dQ_{Y,\epsilon}(y|A = 0, W = w)\right) \\ &= \bar{Q}_{b,0}(w) + \epsilon\Big(E_0\left[YS_Y(Y|1, W)|A = 1, w\right] - E_0\left[YS_Y(Y|0, W)|A = 0, w\right]\Big) \\ &\triangleq \bar{Q}_{b,0}(w) + \epsilon h(w),\end{aligned} \tag{2.18}$$

where we note that $\sup_w |h(w)| < \infty$ because Y and S_Y are uniformly bounded.

Pathwise differentiable if (2.3).
Suppose (2.3). Let $B_1 \triangleq \{w : \bar{Q}_{b,0}(w) = 0\}$ and $B_2 \triangleq \{w : \bar{Q}_{b,0}(w) = 0, \max_a \sigma_0(a, w) = 0\}$. Noting that $B_2 \subseteq B_1$ shows

$$\begin{aligned}&E_{P_\epsilon}\left[\left(I(\bar{Q}_{b,\epsilon} > 0) - I(\bar{Q}_{b,0} > 0)\right)\bar{Q}_{b,\epsilon}\right] \\ &= \int_{\mathcal{W}\backslash B_1}\left(I(\bar{Q}_{b,\epsilon} > 0) - I(\bar{Q}_{b,0} > 0)\right)\bar{Q}_{b,\epsilon}dQ_{W,\epsilon} \\ &\quad + \int_{B_1\backslash B_2}\left(I(\bar{Q}_{b,\epsilon} > 0) - I(\bar{Q}_{b,0} > 0)\right)\bar{Q}_{b,\epsilon}dQ_{W,\epsilon} \\ &\quad + \int_{B_2}\left(I(\bar{Q}_{b,\epsilon} > 0) - I(\bar{Q}_{b,0} > 0)\right)\bar{Q}_{b,\epsilon}dQ_{W,\epsilon}.\end{aligned} \tag{2.19}$$

Because $\bar{Q}_{b,0} \neq 0$ on $\mathcal{W}\backslash B_2$, the first term above is $o(|\epsilon|)$ by a slight generalization of Lemma 2 in van der Laan and Luedtke (2014b) to finite measures (since $\Pr_0(\mathcal{W}\backslash B_2)$ may be less than 1). The second term is zero because $\Pr_0(B_1\backslash B_2) = 0$ by (2.3). Let $f(a, w) \triangleq$

$E_0\left[YS_Y(Y|1,W)|A=1,W=w\right]$. For the third term, note that, for $(a,w)\in\{0,1\}\times B_2$,

$$\int_{B_2}\left(I(\bar{Q}_{b,\epsilon}>0)-I(\bar{Q}_{b,0}>0)\right)\bar{Q}_{b,\epsilon}dQ_{W,\epsilon}$$
$$=\epsilon\int_{B_2}\left(I(\bar{Q}_{b,\epsilon}>0)-I(\bar{Q}_{b,0}>0)\right)\left(f(1,w)-f(0,w)\right)dQ_{W,\epsilon}$$

Note that $f(a,w)=\mathrm{Cov}_{P_0}(Y,S_Y(Y|A,W)|A=a,W=w)$ for $a=0,1$ because $E[S_Y|A,W]$ is equal to zero, and thus $f(a,w)=0$ for $(a,w)\in\{0,1\}\times B_2$ since Y has conditional variance 0 given $A=a$ and $W=w$. This shows that the third term in (2.19) is exactly zero. Hence,

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}E_{P_\epsilon}\left[\left(I(\bar{Q}_{b,\epsilon}>0)-I(\bar{Q}_{b,0}>0)\right)\bar{Q}_{b,\epsilon}\right]=0.$$

Thus Ψ has canonical gradient $D(d_0',P_0)$, i.e. the same canonical gradient as the parameter $\Psi_{d_0'}$. Recall that

$$D(d,P)(O)=\frac{I(A=d(W))}{g(A|W)}(Y-\bar{Q}(A,W))+\bar{Q}(d(W),W)-\Psi_d(P).$$

If (2.3) holds, then either i) $Y=\bar{Q}(A,W)$ or ii) $d_0^*=d_0'$ with P_0 probability 1. Thus $D(d_0^*,P_0)=D(d_0',P_0)$ almost surely if (2.3) holds. It follows that Ψ has canonical gradient $D(d_0^*,P_0)$.

Not pathwise differentiable if not (2.3).
We wish to construct a submodel so that (2.4) holds. Let $S_W(w)=0$ for all w. Without loss of generality, suppose that

$$P_0\left(\bar{Q}_{b,0}(W)=0,\sigma_0(1,W)>0\right)>0. \tag{2.20}$$

Let

$$R(w)\triangleq\frac{\mathrm{Pr}_0(Y\le\bar{Q}_0(1,W)|A=1,W=w)}{\mathrm{Pr}_0(Y>\bar{Q}_0(1,W)|A=1,W=w)},$$

where we let $R(w)=\infty$ when $\mathrm{Pr}_0(Y>\bar{Q}_0(1,W)|A=1,W=w)=0$. Define S_Y as follows:

$$S_Y(y|a,w)\triangleq\begin{cases}\min\{1,R(w)\}, & \text{if } a=1 \text{ and } y>\bar{Q}_0(1,w)\\ -\min\{1,1/R(w)\}, & \text{if } a=1 \text{ and } y\le\bar{Q}_0(1,w)\\ 0, & \text{if } a=0.\end{cases}$$

Above we let $\min\{1,1/R(W)\}=0$ when $R(W)=\infty$ and $\min\{1,1/R(W)\}=1$ when $R(W)=0$. Note that $\sup_{w,a,y}|S_Y(y|a,w)|\le 1$ and $E[S_Y|A=a,W=w]=0$ for all a,w. We define B_+ and B_- as follows:

$$B_+\triangleq B_0\cap\{w:h(w)>0\}$$
$$B_-\triangleq B_0\cap\{w:h(w)<0\},$$

where h is defined in (2.18). By (2.20), $\Pr_0(\bar{Q}_{b,0}(W) = 0, 0 < R(W) < \infty) > 0$, and hence $\Pr_0(B_+) > 0$ and $\Pr_0(B_-) > 0$. Let $m(w) \triangleq (I(\bar{Q}_{b,\epsilon}(w) > 0) - I(\bar{Q}_{b,0}(w) > 0))\bar{Q}_{b,\epsilon}(w)$. Below we omit the dependence of m, $Q_{W,0}$, and h on w. The first term in (2.17) yields the following limit for $\epsilon^{-1}\int_{\mathcal{W}} m\, dQ_{W,0}$ as $\epsilon \downarrow 0$:

$$\begin{aligned}
&\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\int_{B_+} m\, dQ_{W,0} + \lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\int_{B_-} m\, dQ_{W,0} + \lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\int_{\mathcal{W}\backslash(B_+\cup B_-)} m\, dQ_{W,0} \\
&= \lim_{\epsilon\downarrow 0}\int_{B_+} I(\epsilon h > 0)h\, dQ_{W,0} + \lim_{\epsilon\downarrow 0}\int_{B_-} I(\epsilon h > 0)h\, dQ_{W,0} + \lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\int_{\mathcal{W}\backslash(B_+\cup B_-)} m\, dQ_{W,0} \\
&= \int_{B_+} h\, dQ_{W,0} > 0, \qquad (2.21)
\end{aligned}$$

where the integral over B_- is equal to zero because the indicator in m is 0 for all $\epsilon > 0$ and the integral over $\mathcal{W}\backslash(B_+ \cup B_-)$ is $o(|\epsilon|)$ because

$$\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\int_{\mathcal{W}\backslash(B_+\cup B_-)} m\, dQ_{W,0} = \lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\int_{\mathcal{W}\backslash B_0} m\, dQ_{W,0} + \lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\int_{\{w:h=0\}\cap B_0} m\, dQ_{W,0} = 0,$$

where we used that the first term is 0 by a slight generalization of Lemma 2 in van der Laan and Luedtke (2014b) to finite measures and the second term is 0 because $\bar{Q}_{b,\epsilon} = 0$ on $\{w : h = 0\} \cap B_0$. The inequality in (2.21) is strict because $\Pr_0(B_+) > 0$ and $h > 0$ on B_+. Similarly,

$$\lim_{\epsilon\uparrow 0}\frac{1}{\epsilon}\int m\, dQ_{W,0} = \int_{B_-} h\, dQ_{W,0} < 0.$$

Contrasting the above with (2.21) shows that there exists a path about P_0 which results in a fluctuation h for which the limit of the first term in (2.17) divided by ϵ does not exist as $\epsilon \to 0$. But then Ψ cannot be pathwise differentiable: one of the limits in the sum on the right-hand side of (2.17) exists, so the limit on the left-hand side cannot exist. Specifically, suppose c_n has a limit as $n \to \infty$ and $a_n = b_n + c_n$. If b_n does not have a limit, then a_n does not have a limit, since a_n having a limit implies that $b_n = a_n - c_n$ has a limit, contradiction. □

Proofs of results from Section 2.5

Proof of Theorem 2. We have that

$$
\begin{aligned}
&\Gamma_n \left(\hat{\Psi}(P_n) - \Psi(P_0)\right) \\
&= \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}_{n,j}^{-1} \left(\tilde{D}_{n,j}(O_j) - \Psi(P_0)\right) && (2.22) \\
&= \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}_{n,j}^{-1} \left(\left[\tilde{D}_{n,j}(O_j) - \Psi_{d_{n,j}}(P_0)\right] + \left[\Psi_{d_{n,j}}(P_0) - \Psi(P_0)\right]\right) && (2.23) \\
&= \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}_{n,j}^{-1} \left(\tilde{D}_{n,j}(O_j) - \Psi_{d_{n,j}}(P_0)\right) + o_{P_0}(n^{-1/2}) && (2.24) \\
&= \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}_{n,j}^{-1} \left(\tilde{D}_{n,j}(O_j) - E_0\left[\tilde{D}_{n,j}(O_j)|O_1,...,O_{j-1}\right]\right) + R_{1n} + o_{P_0}(n^{-1/2}) && (2.25) \\
&= \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}_{n,j}^{-1} \left(\tilde{D}_{n,j}(O_j) - E_0\left[\tilde{D}_{n,j}(O_j)|O_1,...,O_{j-1}\right]\right) + o_{P_0}(n^{-1/2}). && (2.26)
\end{aligned}
$$

Above (2.22) is a result of moving the $\Psi(P_0)$ into the summation in the defintion of Γ_n, (2.23) adds zero to the line above, (2.24) follows by C5), (2.25) is a consequence of the fact that $\Psi_d(P_0) = P_0\tilde{D}(\bar{Q}, g, d) - E_0[(1 - \frac{g_0(d(W)|W)}{g(d(W)|W)})(\bar{Q}(d(W), W) - \bar{Q}_0(d(W), W))]$ for any fixed $\bar{Q}$, g, and d, and (2.26) follows by C4).

For $j = 1, ..., n - \ell_n$, let

$$
M_{n,j} \triangleq \frac{1}{\sqrt{n-\ell_n}} \frac{\left(\tilde{D}(d_{n,j+\ell_n})(O_{j+\ell_n}) - E_0\left[\tilde{D}(d_{n,j+\ell_n})(O_{j+\ell_n})|O_1,...,O_{j+\ell_n}\right]\right)}{\tilde{\sigma}_{n,j+\ell_n}}.
$$

Note that, for each n, $\{M_{n,j} : j = 1, ..., n - \ell_n\}$ is a discrete-time martingale with respect to the filtration $\mathcal{F}_j$, where each $\mathcal{F}_j$ is the sigma-field generated by $O_1, ..., O_{j+\ell_n}$. In particular, we have that, for all $j \geq 1$, $E_0[M_{n,j}|\mathcal{F}_{j-1}] = 0$. We also have that $\sum_{j=1}^{n-\ell_n} E_0[M_{n,j}^2|\mathcal{F}_{j-1}] = \frac{1}{n-\ell_n}\sum_{j=1}^{n-\ell_n} \frac{\tilde{\sigma}_{0,n,j+\ell_n}^2}{\tilde{\sigma}_{n,j+\ell_n}^2} \to 1$ by C3). Because the conditional Lindeberg condition in C2) holds, the martingale CLT for triangular arrays (see, e.g., Theorem 2 in Gaenssler et al., 1978) shows that

$$
\sum_{j=1}^{n-\ell_n} M_{n,j} \rightsquigarrow N(0,1). \tag{2.27}
$$

Plugging this into (2.26) gives that $\Gamma_n\sqrt{n-\ell_n}\left(\hat{\Psi}(P_n) - \Psi(P_0)\right) \rightsquigarrow N(0,1)$. The asymptotically valid $1-\alpha$ CI is now constructed in the usual way. □

Proof of Corollary 1. In this proof we use "$\lesssim$" to denote less than or equal to up to a positive multiplicative constant. Let $\mathcal{F}_j$ represent the sigma-field generated by $O_1, \ldots, O_j$. Let $\tilde{D}_0 \triangleq \tilde{D}(d_0, \bar{Q}_0, g_0)$ and $s_0^2 \triangleq \mathrm{Var}_{P_0}(\tilde{D}(d_0, \bar{Q}_0, g_0)(O))$. The proof can be broken into four parts, which show that: (1) $\tilde{D}_{n,j}$ approximates $\tilde{D}_0$ in mean-square; (2) $\Gamma_n^{-1} \to s_0$ in probability; (3) $\Gamma_n(\hat{\Psi}(P_n) - \Psi(P_0))$ behaves like an empirical mean of the normalized efficient influence curve; (4) $\hat{\Psi}(P_n)$ is RAL and efficient.

Part 1: $\tilde{D}_{n,j}$ approximates $\tilde{D}_0$. Note that

$$
\begin{aligned}
&\frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} E_0\left[\left(\tilde{D}_{n,j} - \tilde{D}_0\right)^2 \middle| \mathcal{F}_{j-1}\right] \\
&\le \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} E_0\left[\left(\tilde{D}(d_{n,j}, \bar{Q}_{n,j}, g_{n,j}) - \tilde{D}(d_0, \bar{Q}_{n,j}, g_{n,j})\right)^2 \middle| \mathcal{F}_{j-1}\right] \\
&\quad + \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} E_0\left[\left(\tilde{D}(d_0, \bar{Q}_{n,j}, g_{n,j}) - \tilde{D}(d_0, \bar{Q}_{n,j}, g_0)\right)^2 \middle| \mathcal{F}_{j-1}\right] \\
&\quad + \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} E_0\left[\left(\tilde{D}(d_0, \bar{Q}_{n,j}, g_0) - \tilde{D}(d_0, \bar{Q}_0, g_0)\right)^2 \middle| \mathcal{F}_{j-1}\right] \\
&\lesssim \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} E_0\left[(d_{n,j}(W) - d_0(W))^2 \middle| \mathcal{F}_{j-1}\right] \\
&\quad + \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} E_0\left[(g_{n,j}(d(W)|W) - g_0(d(W)|W))^2 \middle| \mathcal{F}_{j-1}\right] \\
&\quad + \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} E_0\left[\left(\bar{Q}_{n,j}(d_0(W), W) - \bar{Q}_0(d_0(W), W)\right)^2 \middle| \mathcal{F}_{j-1}\right] \\
&= o_{P_0}(1) \qquad (2.28)
\end{aligned}
$$

where the constant in the second inequality relies on the bounds on Y, $\bar{Q}_{n,j}$, g_0, and $g_{n,j}$.

Part 2: $\Gamma_n^{-1} \to s_0$ in probability. We have that

$$
\begin{aligned}
(\Gamma_n - s_0^{-1})^2 &\le \left(\frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} \tilde{\sigma}_{n,j}^{-1} s_0^{-1} |\tilde{\sigma}_{n,j} - s_0|\right)^2 \lesssim \left(\frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} |\tilde{\sigma}_{n,j} - s_0|\right)^2 \\
&\lesssim \frac{1}{n-\ell_n} \sum_{j=\ell_n+1}^{n} (\tilde{\sigma}_{n,j} - s_0)^2, \qquad (2.29)
\end{aligned}
$$

where the second inequality on the first line holds by the assumed bounds on $\tilde{\sigma}_{n,j}$ and the final inequality holds by Cauchy-Schwarz. Note that, for any positive real numbers x_1, x_2,

$$(x_1 - x_2)^2 \leq 2|x_1^2 - x_2^2|. \tag{2.30}$$

By the above and Condition C3') , we have that

$$\begin{aligned}\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}(\tilde{\sigma}_{n,j}-\tilde{\sigma}_{0,n,j})^2 &\lesssim \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\left|\tilde{\sigma}_{n,j}^2-\tilde{\sigma}_{0,n,j}^2\right| \\ &\lesssim \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\left|\frac{\tilde{\sigma}_{0,n,j}^2}{\tilde{\sigma}_{n,j}^2}-1\right| = o_{P_0}(1).\end{aligned}$$

We also have that

$$\begin{aligned}\frac{1}{n-\ell_n}\sum_{j=1}^{n}(\tilde{\sigma}_{0,n,j}-s_0)^2 &\leq \frac{2}{n-\ell_n}\sum_{j=1}^{n}|\tilde{\sigma}_{0,n,j}^2-s_0^2| \\ &= \frac{2}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\left|E_0\left[\tilde{D}_{n,j}^2-\tilde{D}_0^2|\mathcal{F}_{j-1}\right]+E_0\left[\tilde{D}_{n,j}|\mathcal{F}_{j-1}\right]^2-E_0\left[\tilde{D}_0|\mathcal{F}_{j-1}\right]^2\right| \\ &\lesssim \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}E_0\left[\left|\tilde{D}_{n,j}-\tilde{D}_0\right||\mathcal{F}_{j-1}\right] \\ &\lesssim \sqrt{\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}E_0\left[\left(\tilde{D}_{n,j}-\tilde{D}_0\right)^2|\mathcal{F}_{j-1}\right]},\end{aligned}$$

where: the first inequality holds by (2.30); the equality holds by the definition of conditional variance; the second inequality holds by twice using that $x_1^2 - x_2^2 = (x_1 + x_2)(x_1 - x_2)$, the strong positivity assumption, and the bounds on Y and $\bar{Q}_{n,j}$; and the final inequality holds by the Cauchy-Schwarz inequality applied to the expectations and the concavity of $x \mapsto \sqrt{x}$. By (2.28), the upper bound above is $o_{P_0}(1)$. By the triangle inequality and the previous two indented equations,

$$\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}(\tilde{\sigma}_{n,j}-s_0)^2 \leq \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\left[(\tilde{\sigma}_{n,j}-\tilde{\sigma}_{0,n,j})^2+(\tilde{\sigma}_{0,n,j}-s_0)^2\right] = o_{P_0}(1). \tag{2.31}$$

Plugging this into (2.29) shows that $\Gamma_n = s_0^{-1} + o_{P_0}(1)$. By the continuous mapping theorem, $\Gamma_n^{-1} = s_0 + o_{P_0}(1)$.

Part 3: $\Gamma_n(\hat{\Psi}(P_n) - \Psi(P_0))$ behaves like an empirical mean. For each $n > 1$ and $j = \ell_n + 1, ..., n$, define

$$M'_{n,j} \triangleq \frac{\tilde{D}_{n,j}(O_j) - E_0\left[\tilde{D}_{n,j}(O)|\mathcal{F}_{j-1}\right]}{\tilde{\sigma}_{n,j}} - \frac{\tilde{D}_0(O_j) - E_0\left[\tilde{D}_0(O) \mid \mathcal{F}_{j-1}\right]}{s_0}.$$

We first show that $\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^n M'_{n,j} \to 0$ in probability. Note that

$$\begin{aligned}
V'_{n,j} \triangleq \mathrm{Var}_{P_0}\left(M'_{n,j}\middle|\, \mathcal{F}_{j-1}\right) &= E_0\left[\left(\frac{\tilde{D}_{n,j}(O_j)}{\tilde{\sigma}_{n,j}} - \frac{\tilde{D}_0(O_j)}{s_0}\right)^2\middle|\, \mathcal{F}_{j-1}\right] \\
&\le E_0\left[\left(\frac{\tilde{D}_{n,j}(O_j)}{\tilde{\sigma}_{n,j}} - \frac{\tilde{D}_0(O_j)}{\tilde{\sigma}_{n,j}}\right)^2\middle|\, \mathcal{F}_{j-1}\right] + E_0\left[\left(\frac{\tilde{D}_0(O_j)}{\tilde{\sigma}_{n,j}} - \frac{\tilde{D}_0(O_j)}{s_0}\right)^2\middle|\, \mathcal{F}_{j-1}\right] \\
&\lesssim E_0\left[\left(\tilde{D}_{n,j}(O_j) - \tilde{D}_0(O_j)\right)^2\middle|\, \mathcal{F}_{j-1}\right] + E_0\left[(\tilde{\sigma}_{n,j} - s_0)^2\middle|\, \mathcal{F}_{j-1}\right]
\end{aligned}$$

where the constants in the second inequality rely on the bounds on $g_{n,j}$, g_0, $\bar{Q}_{n,j}$, Y, $\tilde{\sigma}_{0,n,j}$, and s_0. By (2.28) and (2.31),

$$\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^n V'_{n,j} = o_{P_0}(1). \tag{2.32}$$

Fix $\epsilon, \delta > 0$ and let $v_{\epsilon,\delta} \triangleq \frac{\epsilon^2}{\log(4/\delta)}$. We will show that there exists some N such that

$$\Pr_0\left(\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^n M'_{n,j} \ge \epsilon\right) < \delta \text{ for all } n \ge N. \tag{2.33}$$

Note that

$$\begin{aligned}
&\Pr_0\left(\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^n M'_{n,j} \ge \epsilon\right) \\
&= \Pr_0\left(\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^n M'_{n,j} \ge \epsilon, \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^n V'_{n,j} \le v_{\epsilon,\delta}\right) \\
&\quad + \Pr_0\left(\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^n M'_{n,j} \ge \epsilon, \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^n V'_{n,j} > v_{\epsilon,\delta}\right).
\end{aligned}$$

We will bound the terms on the right separately. By our bounding assumptions, there exists some $m^* \in (0,\infty)$ such that $\Pr_0(\sup_{j\le n}|M_{n,j}| < m^*) = 1$. By Bernstein's inequality for martingale difference sequences with bounded increments (see, e.g, Steiger, 1969; Theorem

1.6 of Freedman, 1975), we have that

$$\begin{aligned}
&\Pr_0\left(\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^{n} M'_{n,j} \geq \epsilon, \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} V'_{n,j} \leq v_{\epsilon,\delta}\right) \\
&\leq \Pr_0\left(\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^{\tilde{n}} M'_{n,j} \geq \epsilon, \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{\tilde{n}} V'_{n,j} \leq v_{\epsilon,\delta} \text{ for some } \tilde{n} \in \{\ell_n+1, ..., n\}\right) \\
&\leq \Pr_0\left(\sum_{j=\ell_n+1}^{\tilde{n}} \frac{M'_{n,j}}{m^*} \geq \frac{\epsilon\sqrt{n-\ell_n}}{m^*}, \sum_{j=\ell_n+1}^{\tilde{n}} \frac{V'_{n,j}}{(m^*)^2} \leq \frac{v_{\epsilon,\delta}(n-\ell_n)}{(m^*)^2} \text{ for some } \tilde{n} \in \{\ell_n+1, ..., n\}\right) \\
&\leq \exp\left(-\frac{\epsilon^2\sqrt{n-\ell_n}}{2(m^*\epsilon + v_{\epsilon,\delta}\sqrt{n-\ell_n})}\right) \overset{n\to\infty}{\longrightarrow} \delta/4.
\end{aligned}$$

It follows that there exists some N_1 such that the upper bound above is less than or equal to $\delta/2$ for all $n \geq N_1$. We also have that

$$\Pr_0\left(\frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^{n} M'_{n,j} \geq \epsilon, \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} V'_{n,j} > v_{\epsilon,\delta}\right) \leq \Pr_0\left(\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} V'_{n,j} \geq v_{\epsilon,\delta}\right).$$

By (2.32), there exists some N_2 so that the upper bound above is no greater than $\delta/4$ for all $n \geq N_2$. Combining the previous two sets of inequalities shows that (2.33) is satisfied for $N \triangleq \max\{N_1, N_2\}$. Thus $\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} M'_{n,j} = o_{P_0}(\sqrt{n-\ell_n})$. Because $\ell_n = o(n)$, $\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} M'_{n,j} = o_{P_0}(n^{-1/2})$. Combining this with (2.26) shows that

$$\begin{aligned}
\Gamma_n\left(\hat{\Psi}(P_n) - \Psi(P_0)\right) &= \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} \frac{\tilde{D}_{n,j}(O_j) - E_0[\tilde{D}_{n,j}(O)|\mathcal{F}_{j-1}]}{\tilde{\sigma}_{n,j}} + o_{P_0}(n^{-1/2}) \\
&= s_0^{-1}\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\left(\tilde{D}_0(O_j) - E_0\left[\tilde{D}_0(O)\right]\right) \\
&\quad + \frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n} M'_{n,j} + o_{P_0}(n^{-1/2}) \\
&= s_0^{-1}\frac{1}{n-\ell_n}\sum_{j=\ell_n+1}^{n}\left(\tilde{D}_0(O_j) - E_0\left[\tilde{D}_0(O)\right]\right) + o_{P_0}(n^{-1/2}) \\
&= s_0^{-1}\frac{1}{n}\sum_{j=1}^{n}\left(\tilde{D}_0(O_j) - E_0\left[\tilde{D}_0(O)\right]\right) + o_{P_0}(n^{-1/2}),
\end{aligned}$$

where the final equality uses that $\ell_n = o(n)$ and that $\tilde{D}_0$ is bounded.

Part 4: $\hat{\Psi}(P_n)$ is RAL and efficient. Combining Parts 2 and 3 shows that

$$\hat{\Psi}(P_n) - \Psi(P_0) = \Gamma_n^{-1}\Gamma_n\left(\hat{\Psi}(P_n) - \Psi(P_0)\right) = (s_0 + o_{P_0}(1))\Gamma_n\left(\hat{\Psi}(P_n) - \Psi(P_0)\right)$$
$$= \frac{1}{n}\sum_{j=1}^{n}\left(\tilde{D}_0(O_j) - E_0\left[\tilde{D}_0(O)\right]\right) + o_{P_0}(n^{-1/2}).$$

Thus $\hat{\Psi}(P_n)$ is an asymptotically linear estimator of $\Psi(P_0)$ with influence curve $D(d_0, P_0) = \tilde{D}_0(O_j) - E_0\left[\tilde{D}_0(O)\right]$. If P_0 satisfies (2.3) so that $D(d_0, P_0) = D(d_0^*, P_0)$ almost surely, then Theorem 1 shows that $D(d_0^*, P_0)$ is the efficient influence curve of Ψ. By Proposition 1 of Section 3.3 in Bickel et al. (1993), it follows that (2.3) holds if and only if $\hat{\Psi}(P_n)$ is a RAL estimator and is asymptotically efficient among all RAL estimators. □

Proof of Theorem 3. The below is an abbreviated version of (2.22) through (2.26) and (2.27), with an added inequality which holds because $R_{2n} \leq 0$:

$$\sqrt{n-\ell_n}\Gamma_n\left(\hat{\Psi}(P_n) - \Psi(P_0)\right)$$
$$= \frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^{n}\tilde{\sigma}_{n,j}^{-1}\left(\left[\tilde{D}_{n,j}(O_j) - \Psi_{d_{n,j}}(P_0)\right] + \left[\Psi_{d_{n,j}}(P_0) - \Psi(P_0)\right]\right)$$
$$\leq \frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^{n}\tilde{\sigma}_{n,j}^{-1}\left(\tilde{D}_{n,j}(O_j) - \Psi_{d_{n,j}}(P_0)\right)$$
$$= \frac{1}{\sqrt{n-\ell_n}}\sum_{j=\ell_n+1}^{n}\tilde{\sigma}_{n,j}^{-1}\left(\tilde{D}_{n,j}(O_j) - E_0\left[\tilde{D}_{n,j}(O_j)|O_1,...,O_{j-1}\right]\right) + o_{P_0}(1),$$

which converges to a standard normal by the central limit theorem. Thus,

$$\liminf_{n\to\infty}\Pr_0\left(\sqrt{n-\ell_n}\Gamma_n\left(\hat{\Psi}(P_n) - \Psi(P_0)\right) \leq z_{1-\alpha}\right) \geq 1-\alpha.$$

The first result follows by rearranging terms in the probability statement. The second result is an immediate corollary of Theorem 2. □

Proofs of results from Section 2.7

Proof of Lemma 1. By the almost sure representation theorem (see, e.g., Theorem 1.10.3 in Billingsley, 1999), there exists a probability space $(\Omega', \mathcal{F}', P')$ and a sequence of random variables $R_n' : \Omega' \to \mathbb{R}$ such that $n^\beta R_n' \stackrel{d}{=} n^\beta R_n$ and $n^\beta R_n'(\omega') \to 0$ for all $\omega' \in \Omega'$. Fix $\epsilon > 0$ and $\omega' \in \Omega'$. There exists some N that, for all $n \geq N$, $n^\beta|R_n'(\omega')| < \frac{(1-\beta)\epsilon}{2}$. Also note that

$$\frac{1}{n^{1-\beta}}\sum_{j=1}^{n} j^{-\beta} \leq \frac{1}{n^{1-\beta}}\int_1^n (j-1)^{-\beta}dj = \frac{1}{1-\beta}.$$

Hence, for all $n \geq N$,

$$\begin{aligned}\frac{1}{n^{1-\beta}}\sum_{j=1}^{n}|R'_j(\omega')| &= \frac{1}{n^{1-\beta}}\sum_{j=1}^{N-1}|R'_j(\omega')| + \frac{1}{n^{1-\beta}}\sum_{j=N}^{n}\frac{1}{j^\beta}j^\beta|R'_j(\omega')| \\ &< \frac{1}{n^{1-\beta}}\sum_{j=1}^{N-1}|R'_j(\omega')| + \frac{(1-\beta)\epsilon}{2n^{1-\beta}}\sum_{j=N}^{n}\frac{1}{j^\beta} \\ &\leq \frac{1}{n^{1-\beta}}\sum_{j=1}^{N-1}|R'_j(\omega')| + \frac{\epsilon}{2}.\end{aligned}$$

It follows that $\frac{1}{n^{1-\beta}}\sum_{j=1}^{n}|R'_j(\omega')| < \epsilon$ for all n large enough, and thus that $\frac{1}{n^{1-\beta}}\sum_{j=1}^{n}R'_j(\omega')$ converges to 0 as $n \to \infty$. Noting that $\frac{1}{n^{1-\beta}}\sum_{j=1}^{n}R_j \overset{d}{=} \frac{1}{n^{1-\beta}}\sum_{j=1}^{n}R'_j(\omega')$ for all n, we have that $\frac{1}{n}\sum_{j=1}^{n}R_j = o_{P_0}(n^{-\beta})$. □

Proof of Theorem 4. Let $\tilde{\mathcal{D}}_1 \triangleq \{\tilde{D}(d,\bar{Q},g) : d,\bar{Q},g\}$, $\tilde{\mathcal{D}}_2 \triangleq \{\tilde{D}^2(d,\bar{Q},g) : d,\bar{Q},g\}$, and $j^* \triangleq \min\{j : \delta_j \leq \delta_0\}$. The class $\tilde{\mathcal{D}}_1$ is P_0 Glivenko-Cantelli (GC) by assumption, and $\tilde{\mathcal{D}}_2$ is GC by Theorem 2 of van der Vaart and Wellner (2000). For all $j \geq j^*$, we have that

$$\begin{aligned}\left|\tilde{\sigma}_j^2 - \tilde{\sigma}_{0,j}^2\right| \leq &\left|\frac{1}{j-1}\sum_{i=1}^{j-1}\tilde{D}_j^2(O_i) - E_0\left[\tilde{D}_j^2(O)\middle| O_1,...,O_{j-1}\right]\right| \\ &+ \left|\left(\frac{1}{j-1}\sum_{k=1}^{j-1}\tilde{D}_j(O_k)\right)^2 - E_0\left[\tilde{D}_j(O)\middle| O_1,...,O_{j-1}\right]^2\right|. \qquad (2.34)\end{aligned}$$

The first term on the right converges to 0 in probability because $\tilde{\mathcal{D}}_2$ is GC. For the second term, the mean value theorem shows that

$$\begin{aligned}&\left(\frac{1}{j-1}\sum_{k=1}^{j-1}\tilde{D}_j(O_k)\right)^2 - E_0\left[\tilde{D}_j(O)\middle| O_1,...,O_{j-1}\right]^2 \\ &= 2m_j\underbrace{\left(\frac{1}{j-1}\sum_{k=1}^{j-1}\tilde{D}_j(O_k) - E_0\left[\tilde{D}_j(O)\middle| O_1,...,O_{j-1}\right]\right)}_{\triangleq \|P_j - P_0\|_{\tilde{\mathcal{D}}_1}},\end{aligned}$$

where m_j is an intermediate value between the two squared values on the first line. Using that $\tilde{\mathcal{D}}_1$ is a GC class, we have that m_j converges to $E_0[\tilde{D}_j(O)|O_1,...,O_{j-1}]$ in probability and $\|P_j - P_0\|_{\tilde{\mathcal{D}}_1} = o_{P_0}(1)$. Thus the above is $o_{P_0}(1)$, and plugging this into (2.34) shows that $|\tilde{\sigma}_j^2 - \tilde{\sigma}_{0,j}^2| = o_{P_0}(1)$. The continuous mapping theorem shows that (2.13) is also satisfied. Combining this with Lemma 1 with $\beta = 0$ shows that Condition C3) is satisfied. □

Proof of Theorem 5. In this proof we will omit the dependence of d_0^*, d_n, $\bar{Q}_{b,0}$, and $\bar{Q}_{b,n}$ on W in the notation. Suppose that $\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0} = o_{P_0}(1)$. This part of the proof mimics the proof of Lemma 5.2 in Audibert and Tsybakov (2007). For any $t > 0$,

$$
\begin{aligned}
&|\Psi_{d_n}(P_0) - \Psi_{d_0^*}(P_0)| \\
&= E_0[|\bar{Q}_{b,0}|I(d_0^* \neq d_n)] \\
&= E_0[|\bar{Q}_{b,0}|I(d_0^* \neq d_n)I(0 < |\bar{Q}_{b,0}| \leq t)] + E_0[|\bar{Q}_{b,0}|I(d_0^* \neq d_n)I(|\bar{Q}_{b,0}| > t)] \\
&\leq E_0[|\bar{Q}_{b,n} - \bar{Q}_{b,0}|I(0 < |\bar{Q}_{b,0}| \leq t)] + E_0[|\bar{Q}_{b,n} - \bar{Q}_{b,0}|I(|\bar{Q}_{b,n} - \bar{Q}_{b,0}| > t)] \\
&\leq \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0} \mathrm{Pr}_0(0 < |\bar{Q}_{b,0}| \leq t)^{1/2} + \frac{\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0}^2}{t} \\
&\leq \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0} C_0^{1/2} t^{\alpha/2} + \frac{\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0}^2}{t},
\end{aligned}
$$

where the first inequality holds because $d_0^* \neq d_n$ implies that $|\bar{Q}_{b,n} - \bar{Q}_{b,0}| > |\bar{Q}_{b,0}|$, the second inequality holds by the Cauchy-Schwarz and Markov inequalities, and the third inequality holds by (2.16). The first result follows by optimizing over t to find that the upper bound is minimized when $t = C\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{2,P_0}^{2(1+\alpha)/(2+\alpha)}$ for a constant C which depends on C_0 and α.

Now suppose that $\left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0} = o_{P_0}(1)$. Note that

$$
\begin{aligned}
|\Psi_{d_n}(P_0) - \Psi_{d_0^*}(P_0)| &= E_0\left|I(d_n \neq d_0^*)\bar{Q}_{b,0}\right| \\
&\leq E_0\left[I(0 < |\bar{Q}_{b,0}| \leq |\bar{Q}_{b,n} - \bar{Q}_{b,0}|)|\bar{Q}_{b,0}|\right] \\
&\leq E_0\left[I\left(0 < |\bar{Q}_{b,0}| \leq \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0}\right)|\bar{Q}_{b,0}|\right] \\
&\leq \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0} \mathrm{Pr}_0\left(0 < |\bar{Q}_{b,0}| \leq \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0}\right).
\end{aligned}
$$

By (2.16), $|\Psi_{d_n}(P_0) - \Psi_{d_0^*}(P_0)| \lesssim \left\|\bar{Q}_{b,n} - \bar{Q}_{b,0}\right\|_{\infty,P_0}^{1+\alpha}$. □

Chapter 3

Individualized Treatments Under Limited Resources

3.1 Introduction

In this chapter, we consider a resource constraint under which there is a maximum proportion of the population that can be treated. Given this constraint, we develop a root-n rate estimator for the optimal R-C value and corresponding confidence intervals. We show that our estimator is efficient among all regular and asymptotically linear estimators in our model $\mathcal{M}$ under conditions. When the baseline covariates are continuous and the resource constraint is active, i.e. when the optimal R-C value is less than the optimal unconstrained value, these conditions are far more reasonable than the non-exceptional law assumption needed for regular estimation of the optimal unconstrained value discussed in Chapter 2.

We now give a brief outline of the chapter. Section 3.2 defines the statistical estimation problem of interest, gives an expression for the optimal deterministic rule under a condition, and gives a general expression for the optimal stochastic rule. Section 3.3 presents our estimator of the optimal R-C value. Section 3.4 presents conditions under which the optimal R-C value is pathwise differentiable, and gives an explicit expression for the canonical gradient under these conditions. Section 3.5 describes the properties of our estimator, including how to develop confidence intervals for the optimal R-C value. Section 3.6 presents our simulation methods. Section 3.7 presents our simulation results. Section 3.8 closes with a discussion and areas of future research. All proofs are given in Section 3.9.

3.2 Optimal resource-constrained rule and value

In this chapter we assume that the outcome Y has support in the closed unit interval. Little generality is lost with the bound on Y, given that any continuous outcome bounded can be rescaled to the unit interval via a linear transformation. Suppose that the treatment resource is limited so that at most a $\kappa \in (0, 1)$ proportion of the population can receive the treatment $A = 1$. A deterministic treatment rule $\tilde{d}$ takes as input a covariate vector $w \in \mathcal{W}$ and outputs a binary treatment decision $\tilde{d}(w)$. The stochastic treatment rules considered in this chapter are maps from $\mathcal{U} \times \mathcal{W}$ to $\{0, 1\}$, where $\mathcal{U}$ is the support of some random variable $U \sim P_U$. If d is a stochastic rule and $u \in \mathcal{U}$ is fixed, then $d(u, \cdot)$ represents a deterministic treatment rule. Throughout this chapter we will let U be drawn independently of all draws from P_0. We will be consistent with the use of the tilde to represent deterministic rules and lack of tilde to represent stochastic rules so that throughout $\tilde{d} : \mathcal{W} \to \mathbb{R}$ and $d : \mathcal{U} \times \mathcal{W} \to \mathbb{R}$.

An optimal R-C deterministic regime at P is defined as a deterministic regime $\tilde{d}$ which solves the optimization problem

$$\text{Maximize } \Psi_{\tilde{d}}(P) \text{ subject to } E_0[\tilde{d}(W)] \leq \kappa. \tag{3.1}$$

We call the optimal value under an R-C deterministic regime $\tilde{\Psi}(P)$. For a stochastic regime d, let $\Psi_d(P) \triangleq E_{P_U}[\Psi_{d(U,\cdot)}(P)]$ represent the value of d. An optimal R-C stochastic regime at P is defined as a stochastic treatment regime d which solves the optimization problem

$$\text{Maximize } \Psi_d(P) \text{ subject to } E_{P_U \times P}[d(U, W)] \leq \kappa. \tag{3.2}$$

We call the optimal value under a R-C stochastic regime $\Psi(P)$. Because any deterministic regime can be written as a stochastic regime which does not rely on the stochastic mechanism U, we have that $\Psi(P) \geq \tilde{\Psi}(P)$.

Let S_P represent the survival function of the blip function $\bar{Q}_{b,P}$, i.e. $\tau \mapsto \Pr_P(\bar{Q}_{b,P}(W) > \tau)$. Let

$$\begin{aligned} \eta_P &\triangleq \inf\{\tau : S_P(\tau) \leq \kappa\} \\ \tau_P &\triangleq \max\{\eta_P, 0\}. \end{aligned} \tag{3.3}$$

For notational convenience we let $S_0 \triangleq S_{P_0}$, $\eta_0 \triangleq \eta_{P_0}$, and $\tau_0 \triangleq \tau_{P_0}$.

Define the deterministic treatment rule $\tilde{d}_P$ as $w \mapsto I(\bar{Q}_{b,P}(w) > \tau_P)$, and for notational convenience let $\tilde{d}_0 \triangleq \tilde{d}_{P_0}$. We have the following result.

Theorem 6. *If $\Pr_P(\bar{Q}_{b,P}(W) = \tau_P) = 0$, then the $\tilde{d}_P$ is an optimal deterministic rule satisfying the resource constraint, i.e. $\Psi_{\tilde{d}_P}(P)$ attains the maximum described in (3.1).*

One can in fact show that $\tilde{d}_P$ is the P almost surely unique optimal deterministic regime under the stated condition. We do not treat the case where $\Pr_P(\bar{Q}_{b,P}(W) = \tau_P) > 0$ for deterministic regimes, since in this case (3.1) is a more challenging problem: for discrete W with positive treatment effect in all strata, (3.1) is a special case of the $0-1$ knapsack problem, which is NP-hard, though is considered one of the easier problems in this class (Karp, 1972; Korte and Vygen, 2012). In the knapsack problem, one has a collection of items, each with a value and a weight. Given a knapsack that can only carry a limited weight, the objective is to choose which items to bring so as to maximize the value of the items in the knapsack while respecting the weight restriction. Considering the optimization problem over stochastic rather than deterministic regimes yields a fractional knapsack problem, which is known to be solvable in polynomial time (Dantzig, 1957; Korte and Vygen, 2012). The fractional knapsack problem differs from the $0-1$ knapsack problem in that one can pack partial items, with the value of the partial items proportional to the fraction of the item packed.

Define the stochastic treatment rule d_P by its distribution with respect to a random variable drawn from P_U:

$$\Pr_{P_U}\left(d_P(U,w)=1\right)=\begin{cases}\frac{\kappa-S_P(\tau_P)}{\Pr_P(\bar{Q}_{b,P}(W)=\tau_P)}, & \text{if } \bar{Q}_{b,P}(w)=\tau_P \text{ and } \tau_P>0\\ I(\bar{Q}_{b,P}(w)>\tau_P), & \text{otherwise.}\end{cases}$$

If $\Pr_P(\bar{Q}_{b,P}(W)=\tau_P)=0$, then the first case occurs with probability zero, and this the division by this quantity will not prove problematic. We will let $d_0\triangleq d_{P_0}$. Note that $\tilde{d}_P(W)$ and $d_P(U,W)$ are $P_U\times P$ almost surely equal if $\Pr_P(\bar{Q}_{b,P}(W)=\tau_P)=0$ or if $\tau_P\leq 0$, and thus have the same value in these settings. It is easy to show that

$$E_{P_U\times P}[d_P(U,W)]=\kappa \text{ if } \tau_P>0. \tag{3.4}$$

The following theorem establishes the optimality of the stochastic rule d_P in a resource-limited setting.

Theorem 7. *The maximum in (3.2) is attained at $d=d_P$, i.e. d_P is an optimal stochastic rule.*

Note that the above theorem does not claim that d_P is the unique optimal stochastic regime. For discrete W, the above theorem is an immediate consequence of the discussion of the knapsack problem in Dantzig (1957).

In this chapter we focus on the value of the optimal stochastic rule. Nonetheless, the techniques that we present in this chapter will only yield valid inference in the case where the data are generated according to a distribution P_0 for which $\Pr_0(\bar{Q}_{b,0}(W)=\tau_0)=0$. This is analogous to assuming a non-exceptional law in settings where resources are not limited (Robins, 2004; Luedtke and van der Laan, 2014b), though we note that for continuous covariates W this assumption is much more likely if $\tau_0>0$. It seems unlikely that the treatment effect in some positive probability stratum of covariates will concentrate on some arbitrary (determined by the constraint κ) value τ_0. Nonetheless, one could deal with situations where $\Pr_0(\bar{Q}_{b,0}(W)=\tau_0)>0$ using similar martingale-based online estimation techniques to those presented in Chapter 2.

3.3 Estimating the optimal resource-constrained value

We now present an estimation strategy for the optimal R-C rule. The upcoming sections justify this strategy and suggest that it will perform well for a wide variety of data generating distributions. The estimation strategy proceeds as follows:

1. Obtain estimates $\bar{Q}_n$, $\bar{Q}_{b,n}$, and g_n of $\bar{Q}_0$, $\bar{Q}_{b,0}$, and g_0 using any desired estimation strategy which respects the fact that Y is bounded in the unit interval.

2. Estimate the marginal distribution of W with the corresponding empirical distribution.

3. Estimate S_0 with the plug-in estimator S_n given by $\tau \mapsto \frac{1}{n}\sum_{i=1}^{n} I\left(\bar{Q}_{b,n}(w_i) > \tau\right)$.

4. Estimate η_0 with the plug-in estimator $\eta_n \triangleq \inf\{\tau : S_n(\tau) \leq \kappa\}$.

5. Estimate τ_0 with the plug-in estimator given by $\tau_n \triangleq \max\{\eta_n, 0\}$.

6. Estimate d_0 with the plug-in estimator d_n with distribution

$$\Pr_{P_U}(d_n(U,w) = 1) = \begin{cases} \frac{\kappa - S_n(\tau_n)}{\Pr_{P_n}(\bar{Q}_{b,n}(W) = \tau_n)}, & \text{if } \bar{Q}_{b,n}(w) = \tau_n \text{ and } \tau_n > 0 \\ I(\bar{Q}_{b,n}(w) > \tau_n), & \text{otherwise.} \end{cases}$$

7. Run a TMLE for the parameter $\Psi_{d_n}(P_0)$:

   a) For $\tilde{a} \in \{0,1\}$, define $H(a,w) \triangleq \frac{\Pr_{P_U}(d_n(U,w)=a)}{g_n(a|w)}$. Run a univariate logistic regression using:

$$\begin{aligned} \text{Outcome: } & (y_i : i = 1, ..., n) \\ \text{Offset: } & \left(\operatorname{logit} \bar{Q}_n(a_i, w_i) : i = 1, ..., n\right) \\ \text{Covariate: } & (H(a_i, w_i) : i = 1, ..., n) \,. \end{aligned}$$

   Let ϵ_n represent the estimate of the coefficient for the covariate, i.e.

$$\epsilon_n \triangleq \underset{\epsilon \in \mathbb{R}}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} \left[\bar{Q}_n^\epsilon(a_i, w_i) \log y_i + \left(1 - \bar{Q}_n^\epsilon(a_i, w_i)\right) \log(1 - y_i)\right],$$

   where $\bar{Q}_n^\epsilon(a,w) \triangleq \operatorname{logit}^{-1}\left(\operatorname{logit} \bar{Q}_n(a,w) + \epsilon H(a,w)\right)$.

   b) Define $\bar{Q}_n^* \triangleq \bar{Q}_n^{\epsilon_n}$.

   c) Estimate $\Psi_{d_n}(P_0)$ using the plug-in estimator given by

$$\Psi_{d_n}(P_n^*) \triangleq \frac{1}{n} \sum_{i=1}^{n} \sum_{a=0}^{1} \bar{Q}_n^*(a, w_i) \Pr_{P_U}(d_n(U, w_i) = a).$$

We use $\Psi_{d_n}(P_n^*)$ as our estimate of $\Psi(P_0)$. We will denote this estimator $\hat{\Psi}$, where we have defined $\hat{\Psi}$ so that $\hat{\Psi}(P_n) = \Psi_{d_n}(P_n^*)$. Note that we have used a TMLE for the data dependent parameter $\Psi_{d_n}(P_0)$, which represents the value under a *stochastic* intervention d_n. Nonetheless, we assume that $\Pr_{P_0}(\bar{Q}_{b,0}(W) = \tau_0) = 0$ for many of the results pertaining to our estimator $\hat{\Psi}$, i.e. we assume that the optimal R-C rule is deterministic. We view estimating the value under a stochastic rather than deterministic intervention as worthwhile because one can give conditions under which the above estimator is (root-n) consistent for $\Psi(P_0)$ at all laws P_0, even if non-negligible bias invalidates standard Wald-type confidence intervals for the parameter of interest at laws P_0 for which $\Pr_{P_0}(\bar{Q}_{b,0}(W) = \tau_0) > 0$.

We will use P_n^* to denote any distribution for which $\bar{Q}_{P_n^*} = \bar{Q}_n^*$, $g_{P_n^*} = g_n$, and P_n^* has the marginal empirical distribution of W for the marginal distribution of W. We note that such a distribution P_n^* exists provided that $\bar{Q}_n^*$ and g_n fall in the parameter spaces of $P \mapsto \bar{Q}_P(W)$ and $P \mapsto g_P$, respectively.

In practice we recommend estimating $\bar{Q}_0$ and $\bar{Q}_{b,0}$ using an ensemble method such as super-learning to make an optimal bias-variance trade-off (or, more generally, minimize cross-validated risk) between a mix of parametric models and data adaptive regression algorithms (van der Laan et al., 2007; Luedtke and van der Laan, 2014a). If the treatment mechanism g_0 is unknown then we recommend using similar data adaptive approaches to obtain the estimate g_n. If g_0 is known (as in a randomized controlled trial without missingness), then one can either take $g_n = g_0$ or estimate g_0 using a correctly specified parametric model, which we expect to increase the efficiency of estimators when the $\bar{Q}_0$ part of the likelihood is misspecified (van der Laan and Robins, 2003; van der Laan and Luedtke, 2014a).

There is typically little downside to using data adaptive approaches to estimate the needed portions of the likelihood, though we do give a formal empirical process condition in Section 3.5 which describes exactly how data adaptive these estimators can be. If one is concerned about the data adaptivity of the estimators of the needed portions of the likelihood, then one can consider a cross-validated TMLE approach such as that presented in van der Laan and Luedtke (2014a). This approach makes no restrictions on the data adaptivity of the estimators of $\bar{Q}_0$, $\bar{Q}_{b,0}$, or g_0.

We now outline the main results of this chapter, which hold under appropriate consistency and regularity conditions.

- Asymptotic linearity of $\hat{\Psi}$:

$$\hat{\Psi}(P_n) - \Psi(P_0) = \frac{1}{n}\sum_{i=1}^{n} D_0(O_i) + o_{P_0}(n^{-1/2}),$$

  with D_0 a known function of P_0.

- $\hat{\Psi}$ is an asymptotically efficient estimate of $\Psi(P_0)$.

- One can obtain a consistent estimate σ_n^2 for the variance of $D_0(O)$. An asymptotically valid 95% confidence intervals for $\Psi(P_0)$ given by $\hat{\Psi}(P_n) \pm 1.96\sigma_n/\sqrt{n}$.

The upcoming sections give the consistency and regularity conditions which imply the above results.

3.4 Canonical gradient of the optimal resource-constrained value

The pathwise derivative of Ψ will provide a key ingredient for analyzing the asymptotic properties of our estimator. We refer the reader to Pfanzagl (1990) and Bickel et al. (1993) for an overview of the crucial role that the pathwise derivative plays in semiparametric efficiency theory. We remind the reader that an estimator $\hat{\Phi}$ is an asyptotically linear estimator of a parameter $\Phi(P_0)$ with influence curve IC_{P_0} provided that

$$\hat{\Phi}(P_n) - \Phi(P_0) = \frac{1}{n}\sum_{i=1}^{n} IC_{P_0}(O_i) + o_{P_0}(n^{-1/2}).$$

If Φ is pathwise differentiable with canonical gradient IC_{P_0}, then $\hat{\Phi}$ is RAL and asymptotically efficient (minimum variance) among all such RAL estimators of $\Phi(P_0)$ (Pfanzagl, 1990; Bickel et al., 1993).

For $o \in \mathcal{O}$, a deterministic rule $\tilde{d}$, and a real number τ, define

$$D_1(\tilde{d}, P)(o) \triangleq \frac{I(a = \tilde{d}(w))}{g_P(a|w)}\left(y - \bar{Q}_P(a, w)\right)$$
$$D_2(\tilde{d}, P)(o) \triangleq \bar{Q}_P(\tilde{d}(w), w) - E_P\bar{Q}_P(\tilde{d}(W), W),$$

where $g_P(a|W) \triangleq \Pr_P(A = a|W)$. We will let $g_0 \triangleq g_{P_0}$. We note that $D_1(\tilde{d}, P) + D_2(\tilde{d}, P)$ is the efficient influence curve of the parameter $\Psi_{\tilde{d}}(P)$.

Let d be some stochastic rule. The canonical gradient of Ψ_d is given by

$$IC_d(P)(o) \triangleq E_{P_U}[D_1(d(U, w), P)(o) + D_2(d(U, w), P)(o)].$$

Define

$$D(d, \tau, P)(o) \triangleq IC_d(P)(o) - \tau\left(E_{P_U}\left[d(U, w)\right] - \kappa\right).$$

For ease of reference, let $D_0 \triangleq D(d_0, \tau_0, P_0)$. The upcoming theorem makes use of the following assumptions.

C1) g_0 satisfies the positivity assumption: $\Pr_0(0 < g_0(1|W) < 1) = 1$.

C2) $\bar{Q}_{b,0}(W)$ has density f_0 at η_0, and $0 < f_0(\eta_0) < \infty$.

C3) S_0 is continuous in a neighborhood of η_0.

C4) $\Pr_0(\bar{Q}_{b,0}(W) = \tau) = 0$ for all τ in a neighborhood of τ_0.

We now present the canonical gradient of the optimal R-C value.

Theorem 8. *Suppose C1) through C4). Then Ψ is pathwise differentiable at P_0 with canonical gradient D_0.*

Note that C3) implies that $\Pr_0(\bar{Q}_{b,0}(W) = \tau_0) = 0$. Thus d_0 is (almost surely) deterministic and the expectation over P_U in the definition of D_0 is superfluous. Nonetheless, this representation will prove useful when we seek to show that our estimator solves the empirical estimating equation defined by an estimate of $D(d_0, \tau_0, P_0)$.

When the resource constraint is active, i.e. $\tau_0 > 0$, the above theorem shows that Ψ has an additional component over the optimal value parameter when no resource constraints are present (van der Laan and Luedtke, 2014b). The additional component is $\tau_0 \times (E_{P_U}[d_0(U, w)] - \kappa)$, and is the portion of the derivative that relies on the fact that d_0 is estimated and falls on the edge of the parameter space. We note that it is possible that the variance of $D_0(O)$ is greater than the variance of $IC_{d_0}(P_0)(O)$. If $\tau_0 = 0$ then these two variances are the same, so suppose $\tau_0 > 0$. Then, provided that $\Pr_0(\bar{Q}_{b,0}(W) = \tau_0) = 0$, we have that

$$\begin{aligned}&\mathrm{Var}_{P_0}\left(D_0(O)\right) - \mathrm{Var}_{P_0}\left(IC_{d_0}(P_0)\right)\\&= \tau_0\kappa(1-\kappa)\left(\tau_0 - 2E_0\left[\bar{Q}_0(1,W)\middle|\, \tilde{d}_0(W) = 1\right] + 2E_0\left[\bar{Q}_0(0,W)\middle|\, \tilde{d}_0(W) = 0\right]\right).\end{aligned}$$

For any $\kappa \in (0, 1)$, it is possible to exhibit a distribution P_0 which satisfies the conditions of Theorem 8 and for which $\mathrm{Var}_{P_0}(D_0(O)) > \mathrm{Var}_{P_0}(IC_{d_0}(P_0)(O))$. Perhaps more surprisingly, it is also possible to exhibit a distribution P_0 which satisfies the conditions of Theorem 8 and for which $\mathrm{Var}_{P_0}(D_0(O)) < \mathrm{Var}_{P_0}(IC_{d_0}(P_0)(O))$. We omit further the discussion here because the focus of this chapter is on considering the estimating the value from the optimization problem (3.2), rather than discussing how this procedure relates to the estimation of other parameters.

3.5 Results about the proposed estimator

We now show that $\hat{\Psi}$ is an asymptotically linear estimator for $\Psi(P_0)$ with influence curve D_0 provided our estimates of the needed parts of P_0 satisfy consistency and regularity conditions. Our theoretical results are presented in Section 3.5, and the conditions of our main theorem are discussed in Section 3.5.

Inference for $\Psi(P_0)$

For any distributions P and P_0 satisfying positivity, stochastic intervention d, and real number τ, define the following second-order remainder terms:

$$\begin{aligned}R_{10}(d, P) &\triangleq E_{P_U \times P_0}\left[\left(1 - \frac{g_0(d|W)}{g(d|W)}\right)\left(\bar{Q}_P(d, W) - \bar{Q}_0(d, W)\right)\right]\\R_{20}(d) &\triangleq E_{P_U \times P_0}\left[(d - d_0)(\bar{Q}_{b,0}(W) - \tau_0)\right].\end{aligned}$$

Above the reliance of d and d_0 on (U,W) is omitted in the notation. Let $R_0(d,P) \triangleq R_{10}(d,P) + R_{20}(d)$. The upcoming theorem will make use of the following assumptions.

C5) g_0 satisfies the strong positivity assumption: $\Pr_0(\delta < g_0(1|W) < 1-\delta) = 1$ for some $\delta > 0$.

C6) g_n satisfies the strong positivity assumption for a fixed $\delta > 0$ with probability approaching 1: there exists some $\delta > 0$ such that, with probability approaching 1, $\Pr_0(\delta < g_n(1|W) < 1-\delta) = 1$.

C7) $R_0(d_n, P_n^*) = o_{P_0}(n^{-1/2})$.

C8) $E_0\left[(D(d_n,\tau_0,P_n^*)(O) - D_0(O))^2\right] = o_{P_0}(1)$.

C9) $D(d_n,\tau_0,P_n^*)$ belongs to a P_0-Donsker class $\mathcal{D}$ with probability approaching 1.

C10) $\frac{1}{n}\sum_{i=1}^n D(d_n,\tau_0,P_n^*)(O_i) = o_{P_0}(n^{-1/2})$.

We note that the τ_0 in the final condition above only enters the expression in the sum as a multiplicative constant in front of $-E_{P_U}[d(U,w_i)] - \kappa$.

Theorem 9 ($\hat{\Psi}$ is asymptotically linear)**.** *Suppose C2) through C10). Then $\hat{\Psi}$ is a RAL estimator of $\Psi(P_0)$ with influence curve D_0, i.e.*

$$\hat{\Psi}(P_n) - \Psi(P_0) = \frac{1}{n}\sum_{i=1}^n D_0(O_i) + o_{P_0}(n^{-1/2}).$$

Further, $\hat{\Psi}$ is efficient among all such RAL estimators of $\Psi(P_0)$.

Let $\sigma_0^2 \triangleq \mathrm{Var}_{P_0}(D_0)$. By the central limit theorem, $\sqrt{n}\left(\hat{\Psi}(P_n) - \Psi(P_0)\right)$ converges in distribution to a $N(0,\sigma_0^2)$ distribution. Let $\sigma_n^2 \triangleq \frac{1}{n}\sum_{i=1}^n D(d_n,\tau_n,P_n^*)(O_i)^2$ be an estimate of σ_0^2. We now give the following lemma, which gives sufficient conditions for the consistency of τ_n for τ_0.

Lemma 2 (Consistency of τ_n)**.** *Suppose C2) and C3). Also suppose $\bar{Q}_{b,n}$ is consistent for $\bar{Q}_{b,0}$ in $L^1(P_0)$ and that the estimate $\bar{Q}_{b,n}$ belongs to a P_0 Glivenko Cantelli class with probability approaching 1. Then $\tau_n \to \tau_0$ in probability.*

It is easy to verify that conditions similar to those of Theorem 9, combined with the convergence of τ_n to τ_0 as considered in the above lemma, imply that $\sigma_n \to \sigma_0$ in probability. Under these conditions, an asymptotically valid two-sided $1-\alpha$ confidence interval is given by

$$\hat{\Psi}(P_n) \pm z_{1-\alpha/2}\frac{\sigma_n}{\sqrt{n}},$$

where $z_{1-\alpha/2}$ denotes the $1-\alpha/2$ quantile of a $N(0,1)$ random variable.

Discussion of conditions of Theorem 9

Conditions C2) and C3). These are standard conditions used when attempting to estimate the κ-quantile η_0, defined in (3.3). Provided good estimation of $\bar{Q}_{b,0}$, these conditions ensure that gathering a large amount of data will enable one to get a good estimate of the κ-quantile of the random variable $\bar{Q}_{b,0}$. See Lemma 2 for an indication of what is meant by "good estimation" of $\bar{Q}_{b,0}$. It seems reasonable to expect that these conditions will hold when W is continuous and $\eta_0 \neq 0$, since we are assuming that $\bar{Q}_{b,0}$ is not degenerate at the arbitrary (determined by κ) point η_0.

Condition C4). If $\tau_0 > 0$, then C4) is implied by C3). If $\tau_0 = 0$, then C4) is like assuming a non-exceptional law, i.e. that the probability of a there being no treatment effect in a stratum of W is zero. Because τ_0 is not known from the outset, we require something slightly stronger, namely that the probability of *any specific* small treatment effect is zero in a stratum of W is zero. Note that this condition does *not* prohibit the treatment effect from being small, e.g. $\Pr_0(|\bar{Q}_{b,0}(W)| < \tau) > 0$ for all $\tau > 0$, but rather it prohibits there existing a sequence $\tau_m \downarrow 0$ with the property that $\Pr_0(\bar{Q}_{b,0}(W) = \tau_m) > 0$ infinitely often. Thus this condition does not really seem any stronger than assuming a non-exceptional law. If one is concerned about such exceptional laws then we suggest adapting the methods in (Luedtke and van der Laan, 2014b) to the R-C setting.

Condition C5). This condition assumes that people from each stratum of covariates have a reasonable (at least a $\delta > 0$) probability of treatment.

Condition C6). This condition requires that our estimates of g_0 respect the fact that each stratum of covariates has a reasonable probability of treatment.

Condition C7). This condition is satisfied if $R_{10}(d_n, P_n^*) = o_{P_0}(n^{-1/2})$ and $R_{20}(d_n) = o_{P_0}(n^{-1/2})$. The term $R_{10}(d_n, P_n^*)$ takes the form of a typical double robust term that is small if either g_0 or $\bar{Q}_0$ is estimated well, and is second-order, i.e. one might hope that $R_{10}(d_n, P_n^*) = o_{P_0}(n^{-1/2})$, if both both g_0 and $\bar{Q}_0$ are estimated well. One can upper bound this remainder with a product of the $L^2(P_0)$ rates of convergence of these two quantities using the Cauchy-Schwarz inequality. If g_0 is known, then one can take $g_n = g_0$ and this term is zero.

Ensuring that $R_{20}(d_n) = o_{P_0}(n^{-1/2})$ requires a little more work but will still prove to be a reasonable condition. We will use the following margin assumption for some $\alpha > 0$:

$$\Pr_0\left(0 < |\bar{Q}_{b,0} - \tau_0| \leq t\right) \lesssim t^\alpha \text{ for all } t > 0, \tag{3.5}$$

where "$\lesssim$" denotes less than or equal to up to a multiplicative constant. This margin assumption is analogous to that used in Audibert and Tsybakov (2007) and to the condition used in Theorem 5 in the previous chapter. The following result relates the rate of convergence of $R_{20}(d_n)$ to the rate at which $\bar{Q}_{b,n} - \tau_n$ converges to $\bar{Q}_{b,0} - \tau_0$.

Theorem 10. *If (3.5) holds for some* $\alpha > 0$*, then*

i) $|R_{20}(d_n)| \lesssim \left\|(\bar{Q}_{b,n} - \tau_n) - (\bar{Q}_{b,0} - \tau_0)\right\|_{2,P_0}^{2(1+\alpha)/(2+\alpha)}$

ii) $|R_{20}(d_n)| \lesssim \left\|(\bar{Q}_{b,n} - \tau_n) - (\bar{Q}_{b,0} - \tau_0)\right\|_{\infty,P_0}^{1+\alpha}$.

The proof of this lemma is similar to that of Theorem 5 from Chapter 2 so is omitted. Interested readers can find a complete proof of this lemma in Luedtke and van der Laan (2015). If S_0 has a finite derivative at τ_0, as is given by C2), then one can take $\alpha = 1$. The above theorem then implies that $R_{20}(d_n) = o_{P_0}(n^{-1/2})$ if either $\left\|(\bar{Q}_{b,n} - \tau_n) - (\bar{Q}_{b,0} - \tau_0)\right\|_{2,P_0}$ is $o_{P_0}(n^{-3/8})$ or $\left\|(\bar{Q}_{b,n} - \tau_n) - (\bar{Q}_{b,0} - \tau_0)\right\|_{\infty,P_0}$ is $o_{P_0}(n^{-1/4})$.

Condition C8). This is a mild consistency condition which is implied by the $L^2(P_0)$ consistency of d_n, g_n, and $\bar{Q}_n^*$ to d_0, g_0, and $\bar{Q}_0$. We note that the consistency of the intial (unfluctuated) estimate $\bar{Q}_n$ for $\bar{Q}_0$ will imply the consistency of $\bar{Q}_n^*$ to $\bar{Q}_0$ given C6), since in this case $\epsilon_n \to 0$ in probability, and thus $\left\|\bar{Q}_n^* - \bar{Q}_n\right\|_{2,P_0} \to 0$ in probability.

Condition C9). This condition places restrictions on how data adaptive the estimators of d_0, g_0, and $\bar{Q}_0$ can be. We refer the reader to Section 2.10 of van der Vaart and Wellner (1996) for conditions under which the estimates of d_0, g_0, and $\bar{Q}_0$ belonging to Donsker classes implies that $D(d_n, \tau_0, P_n^*)$ belongs to a Donsker class. This condition was avoided for estimating the value function using a cross-validated TMLE in van der Laan and Luedtke (2014a), and using this technique will allow one to avoid the condition here as well.

Condition C10). Using the notation $Pf = \int f(o)dP(o)$ for any distribution P and function $f : \mathcal{O} \to \mathbb{R}$, we have that

$$\begin{aligned} P_n D(d_n, \tau_0, P_n^*) &= P_n D_1(d_n, P_n^*) + P_n D_2(d_n, P_n^*) \\ &\quad - \tau_0 \left(\frac{1}{n} \sum_{i=1}^{n} E_{P_U}[d_n(U, w_i)] - \kappa \right). \end{aligned}$$

The first term is zero by the fluctuation step of the TMLE algorithm and the second term on the right is zero because P_n^* uses the empirical distribution of W for the marginal distribution of W. If $\tau_0 = 0$ then clearly the third term is zero, so suppose $\tau_0 > 0$. Combining (3.4) and the fact that d_n is a substitution estimator shows that the third term is 0 with probability approaching 1 provided that $\tau_n > 0$ with probability approaching 1. This will of course occur if $\tau_n \to \tau_0 > 0$ in probability, for which Lemma 2 gives sufficient conditions.

3.6 Simulation methods

We simulated i.i.d. draws from two data generating distributions at sample sizes 100, 200, and 1000. For each sample size and distribution we considered resource constraints $\kappa = 0.1$

and $\kappa = 0.9$. We ran 2000 Monte Carlo draws of each simulation setting. All simulations were run in `R` (R Core Team, 2014).

We first present the two data generating distributions considered, and then present the estimation strategies used.

Data generating distributions

Simulation 1

For our first data generating distribution, the baseline covariate vector $W = (W_1, ..., W_4)$ is four dimensional and

$$\begin{aligned}
&W_1, W_2, W_3, W_4 \overset{i.i.d.}{\sim} N(0,1) \\
&A|W \sim \text{Bernoulli}(1/2) \\
&\text{logit}\left(E_0\left[Y|A,W,H=0\right]\right) = 1 - W_1^2 + 3W_2 + A\left(5W_3^2 - 4.45\right) \\
&\text{logit}\left(E_0\left[Y|A,W,H=1\right]\right) = -0.5 - W_3 + 2W_1W_2 + A\left(3|W_2| - 1.5\right),
\end{aligned}$$

where H is an unobserved Bernoulli(1/2) variable independent of A, W. For this distribution $E_0[\bar{Q}_0(0,W)] \approx E_0[\bar{Q}_0(1,W)] \approx 0.464$.

We obtained estimates of the approximate optimal R-C optimal value for this data generating distribution using 10^7 Monte Carlo draws. When $\kappa = 0.1$, $\Psi(P_0) \approx 0.511$. When $\kappa = 0.9$, $\Psi(P_0) \approx 0.563$. We note that the resource constraint is not active ($\tau_0 = 0$) when $\kappa = 0.9$.

Simulation 2

For the second data generating distribution, $W \sim \text{Uniform}(-1,1)$, $A|W \sim \text{Bernoulli}(1/2)$, and $Y|A,W \sim \text{Bernoulli}(\bar{Q}_0(A,W))$, where for $\tilde{W} \triangleq W + 5/6$ we define

$$\bar{Q}_0(A,W) - \frac{6}{10} \triangleq \begin{cases} 0, & \text{if } A=1 \text{ and } -1/2 \le W \le 1/3 \\ -\tilde{W}^3 + \tilde{W}^2 - \frac{1}{3}\tilde{W} + \frac{1}{27}, & \text{if } A=1 \text{ and } W < -1/2 \\ -W^3 + W^2 - \frac{1}{3}W + \frac{1}{27}, & \text{if } A=1 \text{ and } W > 1/3 \\ -\frac{3}{10}, & \text{otherwise.} \end{cases}$$

For this distribution $E_0[\bar{Q}_0(0,W)] = 0.3$ and $E_0[\bar{Q}_0(1,W)] \approx 0.583$. This simulation is an example of a case where $\bar{Q}_{b,0}(W) > 0$ almost surely, so any resource constraint will reduce the optimal value from its unconstrained value of 0.583. In particular, we have that $\Psi(P_0) \approx 0.337$ when $\kappa = 0.1$ and $\Psi(P_0) \approx 0.572$ when $\kappa = 0.9$.

Estimating nuisance functions

We treated g_0 as known in both simulations and let $g_n = g_0$. We estimated $\bar{Q}_0$ using the super-learner algorithm with the quasi-log-likelihood loss function (`family=binomial`) and a candidate library of data adaptive (`SL.gam` and `SL.nnet`) and parametric algorithms (`SL.bayesglm`, `SL.glm`, `SL.glm.interaction`, `SL.mean`, `SL.step`, `SL.step.interaction`, and `SL.step.forward`). We refer the reader to Table 2 in the technical report Luedtke and van der Laan (2014a) for a brief description of these algorithms. We estimated $\bar{Q}_{b,0}$ by running a super-learner using the squared error loss function and the same candidate algorithms and used W to predict the outcome $\tilde{Y} \triangleq \frac{2A-1}{g_0(A|W)}(Y-\bar{Y}_n)+\bar{Y}_n$, where $\bar{Y}_n$ represents the sample mean of Y from the n observations. See Luedtke and van der Laan (2014a) for a justification of this estimation scheme.

Once we had our estimates $\bar{Q}_n$, $\bar{Q}_{b,n}$, and g_n we proceeded with the estimation strategy described in Section 3.3.

Evaluating performance

We used three methods to evaluate our proposed approach. First, we looked at the coverage of two-sided 95% confidence intervals for the optimal R-C value. Second, we report the average confidence interval widths. Finally, we looked at the power of the $\alpha = 0.025$ level test $H_0 : \Psi(P_0) = \mu_0$ against $H_1 : \Psi(P_0) > \mu_0$, where $\mu_0 \triangleq E_0[\bar{Q}_0(0, W)]$ is treated as a known quantity. Under causal assumptions, μ_0 can be identified with the counterfactual quantity representing the population mean outcome if, possibly contrary to fact, no one receives treatment. If treatment is not currently being given in the population, one could substitute the population mean outcome (if known) for μ_0. Our test of significance consisted of checking of the lower bound in the two-sided 95% confidence interval is greater than μ_0. If an estimator of $\Psi(P_0)$ is low-powered in testing H_0 against H_1 then clearly the estimator will have little practical value.

3.7 Simulation results

The proposed estimation strategy performed well overall. Figure 3.1 demonstrates the coverage of 95% confidence intervals for the optimal R-C value. Our method performed well in all settings for the highly constrained setting where $\kappa = 0.1$. The results were more mixed for the resource constraint $\kappa = 0.9$. All methods performed well at the largest sample size considered. This supports our theoretical results, which were all asymptotic in nature. For Simulation 1, in which the resource constraint was not active, the coverage dropped off at lower sample sizes. Coverage was somewhat below nominal (80% when $n = 100$ and 84% when $n = 200$) for small sample sizes, but improved when the sample size increased. In Simulation 2, the coverage was better (>91%) for the smaller sample sizes. We note that the resource constraint was still active ($\tau_0 > 0$) when $\kappa = 0.9$ for this simulation, and also that the estimation problem is easier because the baseline covariate was univariate.

Figure 3.2 gives the power of the $\alpha = 0.025$ level test $H_0 : \Psi(P_0) = \mu_0$ against the alternative $H_1 : \Psi(P_0) > \mu_0$. Overall our method appears to have reasonable power in this statistical test. We see that power increases with sample size, the key property of consistent

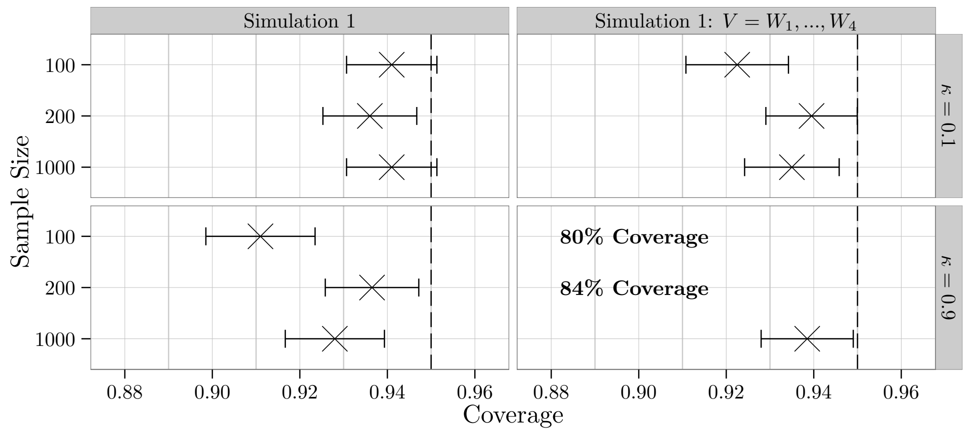

Figure 3.1: Coverage of two-sided 95% confidence intervals. As expected, coverage increases with sample size. The coverage tends to be better for $\kappa = 0.1$ than for $\kappa = 0.9$, though the estimator performed well at the largest sample size (1000) for all simulations and choices of κ. Error bars indicate 95% confidence intervals to account for uncertainty from the finite number of Monte Carlo draws.

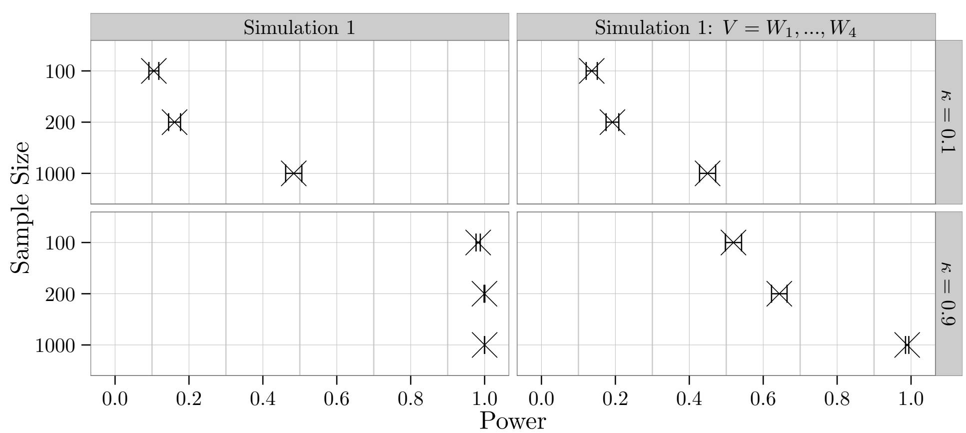

Figure 3.2: Power of the $\alpha = 0.025$ level test of $H_0 : \Psi(P_0) = \mu_0$ against $H_1 : \Psi(P_0) > \mu_0$, where $\mu_0 = E_0[\bar{Q}_0(0, W)]$ is treated as known. Power increases with sample size and κ. Error bars indicate 95% confidence intervals to account for uncertainty from the finite number of Monte Carlo draws.

statistical tests. We also see that power increases with κ, which is unsurprising given that Y is binary and $g_0(a|w)$ is $1/2$ for all a, w. We note that power will not always increase with κ, for example if P_0 is such that $g_0(1|w)$ is very small for individuals with covariate w who are treated at $\kappa = 0.9$ but not at $\kappa = 0.1$. This observation is not meant as a criticism to the estimation scheme that we have presented because we assume that κ will be chosen to reflect real resource constraints, rather than to maximize the power for a test $H_0' : \Psi(P_0) = \mu'$ versus $H_1' : \Psi(P_0) > \mu'$ for some fixed μ'.

We also implemented an estimating equation based estimator for the optimal R-C value and found the two methods performed similarly. We would recommend using the TMLE in practice because it has been shown to be robust to near positivity violations in a wide variety of settings (van der Laan and Rose, 2011). We note that $g_0(1|w) = 1/2$ for all w in both of our simulations, so no near positivity violations occurred. We do not consider the estimation equation approach any further here because the focus of this chapter is on considering the optimization problem (3.2), rather than on comparing different estimation frameworks.

3.8 Discussion and future work

We have considered the problem of estimating the optimal resource-constrained value. Under causal assumptions, this parameter can be identified with the maximum attainable population mean outcome under individualized treatment rules which rely on measured covariates, subject to the constraint that a maximum proportion κ of the population can be treated. We also provided an explicit expression for an optimal stochastic rule under the resource constraint.

We derived the canonical gradient of the optimal R-C value under the key assumption that the treatment effect is not exactly equal to τ_0 in some stratum of covariates which occurs with positive probability. The canonical gradient plays a key role in developing asymptotically linear estimators. We found that the canonical gradient of the optimal R-C value has an additional component when compared to the canonical gradient of the optimal unconstrained value when the resource constraint is active, i.e. when $\tau_0 > 0$.

We presented a targeted minimum loss-based estimator for the optimal R-C value. This estimator was designed to solve the empirical mean of an estimate of the canonical gradient. This quickly yielded conditions under which our estimator is RAL, and efficient among all such RAL estimators. All of these results rely on the condition that the treatment effect is not exactly equal to τ_0 for positive probability stratum of covariates. This assumption is more plausible than the typical non-exceptional law assumption when the covariates are continuous and the constraint is active because it may be unlikely that the treatment effect concentrates on an arbitrary (determined by κ) $\tau_0 > 0$. We note that this pseudo-non-exceptional law assumption has implied that the optimal stochastic rule is almost surely equal to the optimal deterministic rule.

Some resource constraints encountered in practice may not be of the form $E[d(U,W)]$ less than or equal to κ. For example, the cost of distributing the treatment to people may vary based on the values of the covariates. If $c : \mathcal{W} \rightarrow [0,\infty)$ is a cost function, then this constraint may take the form $E_[c(W)d(U,W)] \leq \kappa$. The optimal rule takes the form $(u,w) \mapsto I(\bar{Q}_{b,0}(w) > \tau c(w))$ for w for which $\bar{Q}_{b,0}(w) \neq \tau_0 c(w)$ or $c(w) = 0$, and randomly distributes the remaining resources uniformly among all remaining w. We leave further consideration of this more general resource constraint problem to future work.

Further work is needed to generalize this chapter to the multiple time point setting. Before generalizing the procedure, one must know exactly what form the multiple time point constraint takes. For example, it may be the case that only a κ proportion of the population can be treated at each time point, or it may be the case that treatment can only be administered at a κ proportion of patient-time point pairs. Regardless of which constraint one chooses, it seems that the nice recursive structure encountered in Q-learning may not hold for multiple time point R-C problems. While useful for computational considerations, being able to express the optimal rule using approximate dynamic programming is not necessary for the existence of a good optimal rule estimator, especially when the number of time points is small.

We have not considered the ethical considerations associated with allocating limiting resources to a population. The debate over the appropriate means to distribute limited treatment resources to a population is ongoing (see, e.g., Brock and Wikler, 2009; Macklin and Cowan, 2012; Singh, 2013, for examples in the treatment of HIV/AIDS). Clearly any investigator needs to consider the ethical issues associated with certain resource allocation schemes. Our method is optimal in a particular utilitarian sense (maximizing the expected population mean outcome with respect to an outcome of interest) and yields a treatment strategy which treats individuals who are expected to benefit most from treatment in terms of our outcome of interest. One must be careful to ensure that the outcome of interest truly captures the most important public health implications. Unlike in unconstrained individualized medicine, inappropriately prescribing treatment to a stratum will also have implications for individuals outside of that stratum, namely for the individuals who do not receive treatment due to its lack of availability. We leave further ethical considerations to experts on the matter. It will be interesting to see if there are settings in which it is possible to transform the outcome or add constraints to the optimization problem so that the statistical problem considered in this chapter adheres to the ethical guidelines in those settings.

We have looked to generalize previous works in estimating the value of an optimal individualized treatment regime to the case where the treatment resource is a limited resource, i.e. where it is not possible to treat the entire population. The results in this chapter should allow for the application of optimal personalized treatment strategies to many new problems of interest.

3.9 Proofs

Proofs for Section 3.2

We first state a simple lemma.

Lemma 3. *For a distribution P and a stochastic rule d, we have the following representation for Ψ_d:*

$$\Psi_d(P) \triangleq E_{P_U \times P}\left[d(U,W)\bar{Q}_{b,P}(W)\right] + E_P[\bar{Q}_P(0,W)].$$

Proof of Lemma 3. We have that

$$\begin{aligned}\Psi_d(P) &= E_{P_U\times P}[d(U,W)\bar{Q}_P(1,W)] + E_{P_U\times P}[(1-d(U,W))\bar{Q}_P(0,W)] \\ &= E_{P_U\times P}[d(U,W)(\bar{Q}_P(1,W) - \bar{Q}_P(0,W))] + E_P[\bar{Q}_P(0,W)] \\ &= E_{P_U\times P}[d(U,W)\bar{Q}_{b,P}(W)] + E_P[\bar{Q}_P(0,W)].\end{aligned}$$

□

Proof of Theorem 6. This result will be a consequence of Theorem 7. If $\Pr_P(\bar{Q}_{b,0}(W) = \tau_P) = 0$, then $d_P(U,W)$ is $P_U \times P$ almost surely equal to $\tilde{d}_P(W)$, and thus $\Psi_{\tilde{d}_P}(P) = \Psi_{d_P}(P)$. Thus $(u,w) \mapsto \tilde{d}_P(w)$ is an optimal stochastic regime. Because the class of deterministic regimes is a subset of the class of stochastic regimes, $\tilde{d}_P$ is an optimal deterministic regime. □

Proof of Theorem 7. Let d be some stochastic treatment rule which satisfies the resource constraint. For $(b,c) \in \{0,1\}^2$, define $B_{bc} \triangleq \{(u,w) : d_P(u,w) = b, d(u,w) = c\}$. Note that

$$\begin{aligned}\Psi_{d_P}(P) - \Psi_d(P) &= E_{P_U\times P}\left[(d_P(U,W) - d(U,W))\,\bar{Q}_{b,0}(W)\right] \\ &= E_{P_U\times P}\left[\bar{Q}_{b,0}(W)I((U,W)\in B_{10})\right] - E_{P_U\times P}\left[\bar{Q}_{b,0}(W)I((U,W)\in B_{01})\right] \qquad (3.6)\end{aligned}$$

The $\bar{Q}_{b,0}(W)$ in the first term in (3.6) can be upper bounded by τ_P, and in the second term can be lower bounded by τ_P. Thus,

$$\begin{aligned}\Psi_{d_P}(P) - \Psi_d(P) &\geq \tau_P\left[\Pr_{P_U\times P}((U,W)\in B_{10}) - \Pr_{P_U\times P}((U,W)\in B_{01})\right] \\ &= \tau_P\left[\Pr_{P_U\times P}((U,W)\in B_{10}\cup B_{11}) - \Pr_{P_U\times P}((U,W)\in B_{01}\cup B_{11})\right] \\ &= \tau_P\left(E_{P_U\times P}[d_P(U,W)] - E_{P_U\times P}[d(U,W)]\right).\end{aligned}$$

If $\tau_P = 0$ then the final line is zero. Otherwise, $E_{P_U\times P}[d_P(U,W)] = \kappa$ by (3.4). Because d satisfies the resource constraint, $E_{P_U\times P}[d(U,W)] \leq \kappa$ and thus the final line above is at least zero. Thus $\Psi_{d_P}(P) - \Psi_d(P) \geq 0$ for all τ_P. Because d was arbitrary, d_P is an optimal stochastic rule. □

Proofs for Section 3.4

Proof of Theorem 8. The pathwise derivative of $\Psi(Q)$ is defined as $\frac{d}{d\epsilon}\Psi(Q(\epsilon))\big|_{\epsilon=0}$ along paths $\{P_\epsilon : \epsilon\} \subset \mathcal{M}$. In particular, these paths are chosen so that

$$\begin{aligned} dQ_{W,\epsilon} &= (1+\epsilon H_W(W))dQ_W, \\ &\quad \text{where } EH_W(W)=0 \text{ and } C_W \triangleq \sup_w |H_W(w)| < \infty; \\ dQ_{Y,\epsilon}(Y \mid A,W) &= (1+\epsilon H_Y(Y\mid A,W))dQ_Y(Y\mid A,W), \\ &\quad \text{where } E(H_Y \mid A,W)=0 \text{ and } \sup_{w,a,y}|H_Y(y\mid a,w)| < \infty. \end{aligned}$$

The parameter Ψ is not sensitive to fluctuations of $g_0(a|w) = \Pr_0(a|w)$, and thus we do not need to fluctuate this portion of the likelihood. Let $\bar{Q}_{b,\epsilon} \triangleq \bar{Q}_{b,P_\epsilon}$, $\bar{Q}_\epsilon \triangleq \bar{Q}_{P_\epsilon}$, $d_\epsilon \triangleq d_{P_\epsilon}$, $\eta_\epsilon \triangleq \eta_{P_\epsilon}$, $\tau_\epsilon \triangleq \tau_{P_\epsilon}$, and $S_\epsilon \triangleq S_{P_\epsilon}$. First note that

$$\bar{Q}_{b,\epsilon}(w) = \bar{Q}_{b,0}(w) + \epsilon h_\epsilon(w) \tag{3.7}$$

for an h_ϵ with

$$\sup_{|\epsilon|<1}\sup_w |h_\epsilon(w)| \triangleq C_1 < \infty. \tag{3.8}$$

Note that C4) implies that d_0 is (almost surely) deterministic, i.e. $d_0(U,\cdot)$ is almost surely a fixed function. Let $\tilde{d}$ represent the deterministic rule $w \mapsto I(\bar{Q}_{b,0}(w) > 0)$ to which $d(u,\cdot)$ is (almost surely) equal for all u. By Lemma 3,

$$\begin{aligned} \Psi(P_\epsilon) - \Psi(P_0) &= \int_w \left(E_{P_U}[d_\epsilon(U,W)] - \tilde{d}_0(W)\right)\bar{Q}_{b,\epsilon}dQ_{W,\epsilon} \\ &\quad + \int_w \tilde{d}_0(W)\left(\bar{Q}_{b,\epsilon}dQ_{W,\epsilon} - \bar{Q}_{b,0}dQ_{W,0}\right) + E_{P_\epsilon}\bar{Q}_\epsilon(0,W) - E_0\bar{Q}_0(0,W) \\ &= \int_w \left(E_{P_U}[d_\epsilon(U,W)] - \tilde{d}_0(W)\right)\left(\bar{Q}_{b,\epsilon} - \tau_0\right)dQ_{W,\epsilon} \\ &\quad + \tau_0 \int_w \left(E_{P_U}[d_\epsilon(U,W)]dQ_{W,\epsilon} - \tilde{d}_0(W)dQ_{W,0}\right) \\ &\quad - \tau_0 \int_w \tilde{d}_0(W)\left(dQ_{W,\epsilon} - dQ_{W,0}\right) + \left[\Psi_{d_0}(P_\epsilon) - \Psi_{d_0}(P_0)\right]. \end{aligned} \tag{3.9}$$

Dividing the fourth term by ϵ and taking the limit as $\epsilon \to 0$ gives the pathwise derivative of the mean outcome under the rule that treats d_0 as known. The third term can be written as $-\epsilon\tau_0 \int_w \tilde{d}_0(W)H_W dQ_{W,0}$, and thus the pathwise derivative of this term is $-\int_w \tau_0\tilde{d}_0(W)H_W dQ_{W,0}$. If $\tau_0 > 0$, then $E_{P_U\times P_0}[\tilde{d}_0(W)] = \kappa$. The pathwise derivative of this term is zero if $\tau_0 = 0$. Thus, for all τ_0,

$$\lim_{\epsilon\to 0} -\frac{1}{\epsilon}\tau_0 \int_w \tilde{d}_0(W)\left(dQ_{W,\epsilon} - dQ_{W,0}\right) = \int_w \left(-\tau_0(\tilde{d}_0(w) - \kappa)\right)H_W(w)dQ_{W,0}(w).$$

Thus the third term in (3.9) generates the $w \mapsto -\tau_0(\tilde{d}_0(w) - \kappa)$ portion of the canonical gradient, or equivalently $w \mapsto -\tau_0(E_{P_U}[d_0(U,w)] - \kappa)$. The remainder of this proof is used to show that the first two terms in (3.9) are $o(\epsilon)$.

Step 1: $\eta_\epsilon \to \eta_0$.
We refer the reader to (3.3) for a definition of the quantile $P \mapsto \eta_P$. This is a consequence of the continuity of S_0 in a neighborhood of η_0. For $\gamma > 0$,

$$|\eta_\epsilon - \eta_0| > \gamma \text{ implies that } S_\epsilon(\eta_0 - \gamma) \leq \kappa \text{ or } S_\epsilon(\eta_0 + \gamma) > \kappa. \tag{3.10}$$

For positive constants C_1 and C_W,

$$S_\epsilon(\eta_0 - \gamma) \geq (1 - C_W|\epsilon|) \Pr_0\left(\bar{Q}_{b,\epsilon} > \eta_0 - \gamma\right) \geq (1 - C_W|\epsilon|) S_0(\eta_0 - \gamma + C_1|\epsilon|).$$

Fix $\gamma > 0$ small enough so that S_0 is continuous at $\eta_0 - \gamma$. In this case we have that $S_0(\eta_0 - \gamma + C_1|\epsilon|) \to S_0(\eta_0 - \gamma)$ as $\epsilon \to 0$. By the infimum in the definition of η_0, we know that $S_0(\eta_0 - \gamma) > \kappa$. Thus $S_\epsilon(\eta_0 - \gamma) > \kappa$ for all $|\epsilon|$ small enough.

Similarly, $S_\epsilon(\eta_0 + \gamma) \leq (1 + C_W|\epsilon|) S_0(\eta_0 + \gamma - C_1|\epsilon|)$. Fix $\gamma > 0$ small enough so that S_0 is continuous at $\eta_0 + \gamma$. Then $S_0(\eta_0 + \gamma - C_1|\epsilon|) \to S_0(\eta_0 + \gamma)$ as $\epsilon \to 0$. Condition C2) implies the uniqueness of the κ-quantile of $\bar{Q}_{b,0}$, and thus that $S_0(\eta_0 + \gamma) < \kappa$. It follows that $S_\epsilon(\eta_0 + \gamma) < \kappa$ for all $|\epsilon|$ small enough.

Combining $S_\epsilon(\eta_0 - \gamma) > \kappa$ and $S_\epsilon(\eta_0 + \gamma) < \kappa$ for all ϵ close to zero with (3.10) shows that $\eta_\epsilon \to \eta_0$ as $\epsilon \to 0$.

Step 2: Second term of (3.9) is 0 eventually.
If $\tau_0 = 0$ then the result is immediate, so suppose $\tau_0 > 0$. By the previous step, $\eta_\epsilon \to \eta_0$, which implies that $\tau_\epsilon \to \tau_0 > 0$ by the continuity of the max function. It follows that $\tau_\epsilon > 0$ for ϵ large enough. By (3.4), $\Pr_{P_U \times P_\epsilon}(d_\epsilon(U,W) = 1) = \kappa$ for all sufficiently small $|\epsilon|$ and $\Pr_0(\tilde{d}_0(W) = 1) = \kappa$. Thus the second term of (3.9) is 0 for all $|\epsilon|$ small enough.

Step 3: $\tau_\epsilon - \tau_0 = O(\epsilon)$.
Note that $\kappa < S_\epsilon(\eta_\epsilon - |\epsilon|) \leq (1 + C_W|\epsilon|) S_0(\eta_\epsilon - (1 + C_1)|\epsilon|)$. A Taylor expansion of S_0 about η_0 shows that

$$\begin{aligned}
\kappa &< (1 + C_W|\epsilon|)\left(S_0(\eta_0) + (\eta_\epsilon - \eta_0 - (1 + C_1)|\epsilon|)(-f_0(\eta_0) + o(1))\right) \\
&= \kappa + (\eta_\epsilon - \eta_0 - (1 + C_1)|\epsilon|)(-f_0(\eta_0) + o(1)) + O(\epsilon) \\
&= \kappa - (\eta_\epsilon - \eta_0) f_0(\eta_0) + o(\eta_\epsilon - \eta_0) + O(\epsilon).
\end{aligned} \tag{3.11}$$

The fact that $f_0(\eta_0) \in (0, \infty)$ shows that $\eta_\epsilon - \eta_0$ is bounded above by some $O(\epsilon)$ sequence. Similarly, $\kappa \geq S_\epsilon(\eta_\epsilon + |\epsilon|) \geq (1 - C_W|\epsilon|) S_0(\eta_\epsilon + (1 + C_1)|\epsilon|)$. Hence,

$$\begin{aligned}
\kappa &\geq (1 - C_W|\epsilon|)\left(S_0(\eta_0) + (\eta_\epsilon - \eta_0 + (1 + C_1)|\epsilon|)(-f_0(\eta_0) + o(1))\right) \\
&= \kappa - (\eta_\epsilon - \eta_0) f_0(\eta_0) + o(\eta_\epsilon - \eta_0) + O(\epsilon).
\end{aligned}$$

It follows that $\eta_\epsilon - \eta_0$ is bounded below by some $O(\epsilon)$ sequence. Combining these two bounds shows that $\eta_\epsilon - \eta_0 = O(\epsilon)$, which immediately implies that $\tau_\epsilon - \tau_0 = \max\{O(\epsilon), 0\} = O(\epsilon)$.

Step 4: First term of (3.9) is $o(\epsilon)$.
We know that

$$\bar{Q}_{b,0}(W) - \tau_0 + O(\epsilon) \leq \bar{Q}_{b,\epsilon}(W) - \tau_\epsilon \leq \bar{Q}_{b,0}(W) - \tau_0 + O(\epsilon).$$

By C4), it follows that there exists some $\delta > 0$ such that $\sup_{|\epsilon|<\delta} \Pr_0(\bar{Q}_{b,\epsilon}(W) = \tau_\epsilon) = 0$. By the absolute continuity of $Q_{W,\epsilon}$ with respect to $Q_{W,0}$, $\sup_{|\epsilon|<\delta} \Pr_{P_\epsilon}(\bar{Q}_{b,\epsilon}(W) = \tau_\epsilon) = 0$. It follows that, for all small enough $|\epsilon|$ and almost all u, $d_\epsilon(u,w) = I(\bar{Q}_{b,\epsilon}(w) > \tau_\epsilon)$. Hence,

$$\begin{aligned}
&\int_w \left(E_{P_U}[d_\epsilon(U,W)] - d_0(W)\right)\left(\bar{Q}_{b,\epsilon} - \tau_0\right) dQ_{W,\epsilon} \\
&\quad = \left| \int_w \left(I(\bar{Q}_{b,\epsilon} > \tau_\epsilon) - I(\bar{Q}_{b,0} > \tau_0)\right)\left(\bar{Q}_{b,\epsilon} - \tau_0\right) dQ_{W,\epsilon} \right| \\
&\quad \leq \int_w \left| I(\bar{Q}_{b,\epsilon} > \tau_\epsilon) - I(\bar{Q}_{b,0} > \tau_0)\right| \left(\left|\bar{Q}_{b,0} - \tau_0\right| + C_1|\epsilon|\right) dQ_{W,\epsilon} \\
&\quad \leq \int_w I(|\bar{Q}_{b,0} - \tau_0| \leq |\bar{Q}_{b,0} - \tau_0 - \bar{Q}_{b,\epsilon} + \tau_\epsilon|)\left(\left|\bar{Q}_{b,0} - \tau_0\right| + C_1|\epsilon|\right) dQ_{W,\epsilon} \\
&\quad = \int_w I(0 < |\bar{Q}_{b,0} - \tau_0| \leq |\bar{Q}_{b,0} - \tau_0 - \bar{Q}_{b,\epsilon} + \tau_\epsilon|)\left(\left|\bar{Q}_{b,0} - \tau_0\right| + C_1|\epsilon|\right) dQ_{W,\epsilon} \\
&\quad \leq O(\epsilon) \int_w I(0 < |\bar{Q}_{b,0} - \tau_0| \leq O(\epsilon)) dQ_{W,\epsilon} \\
&\quad \leq O(\epsilon)(1 + C_W|\epsilon|) \Pr_0\left(0 < |\bar{Q}_{b,0} - \tau_0| \leq O(\epsilon)\right),
\end{aligned}$$

where the penultimate inequality holds by Step 3 and (3.7). The last line above is $o(\epsilon)$ because $Pr(0 < X \leq \epsilon) \to 0$ as $\epsilon \to 0$ for any random variable X. Thus dividing the left-hand side above by ϵ and taking the limit as $\epsilon \to 0$ yields zero. □

Proofs for Section 3.5

We give the following lemma before proving Theorem 9.

Lemma 4. *Let P_0 and P be distributions which satisfy the positivity assumption and for which Y is bounded in probability. Let d be some stochastic treatment rule and τ be some real number. We have that* $\Psi_d(P) - \Psi(P_0) = -E_0[D(d,\tau_0,P)(O)] + R_0(d,P)$.

Proof of Lemma 4. Note that

$$\begin{aligned}&\Psi_d(P) - \Psi(P_0) + E_0[D(d, \tau_0, P)(O)] \\ &= \Psi_d(P) - \Psi_d(P_0) + \sum_{j=1}^{2} E_{P_U \times P_0}[D_j(d(U, \cdot), P)(O)] \\ &\quad + \Psi_d(P_0) - \Psi_{d_0}(P_0) - \tau_0 E_{P_U \times P_0}[d(U, W) - \kappa].\end{aligned}$$

Standard calculations show that the first term on the right is equal to $R_{10}(d, P)$ (van der Laan and Robins, 2003). If $\tau_0 > 0$, then (3.4) shows that $\tau_0 E_{P_U \times P_0}[d - \kappa] = \tau_0 E_{P_U \times P_0}[d - d_0]$. If $\tau_0 = 0$, then obviously $\tau_0 E_{P_U \times P_0}[d - \kappa] = \tau_0 E_{P_U \times P_0}[d - d_0]$. Lemma 3 shows that $\Psi_d(P_0) - \Psi_{d_0}(P_0) = E_{P_U \times P_0}[(d - d_0)\bar{Q}_{b,0}]$. Thus the second line above is equal to $R_{20}(d)$. □

Proof of Theorem 9. We make use of empirical process theory notation in this proof so that $Pf = E_P[f(O)]$ for a distribution P and function f. We have that

$$\begin{aligned}&\hat{\Psi}(P_n) - \Psi(P_0) \\ &= -P_0 D(d_n, \tau_0, P_n^*) + R_0(d_n, P_n^*) && \text{(by Lemma 4)} \\ &= (P_n - P_0) D(d_n, \tau_0, P_n^*) + R_0(d_n, P_n^*) + o_{P_0}(n^{-1/2}) && \text{(by Condition C10))} \\ &= (P_n - P_0) D_0 + (P_n - P_0)(D(d_n, \tau_0, P_n^*) - D_0) + R_0(d_n, P_n^*).\end{aligned}$$

The middle term on the last line is $o_{P_0}(n^{-1/2})$ by C5), C6), C8), and C9) (van der Vaart and Wellner, 1996), and the third term is $o_{P_0}(n^{-1/2})$ by C7). This yields the asymptotic linearity result. Proposition 1 in Section 3.3 of Bickel et al. (1993) yields the claim about regularity and asymptotic efficiency when conditions C2), C3), C4), and C5) hold (see Theorem 8). □

Proof of Lemma 2. We will show that $\eta_n \to \eta_0$ in probability, and then the consistency of τ_n follows by the continuous mapping theorem. By C3), there exists an open interval N containing η_0 on which S_0 is continuous. Fix $\eta \in N$. Because $\bar{Q}_{b,n}$ belongs to a Glivenko-Cantelli class with probability approaching 1, we have that

$$\begin{aligned}|S_n(\eta) - S_0(\eta)| &= \left|P_n I(\bar{Q}_{b,n} > \eta) - P_0 I(\bar{Q}_{b,0} > \eta)\right| \\ &\le \left|P_0 \left(I(\bar{Q}_{b,n} > \eta) - I(\bar{Q}_{b,0} > \eta)\right)\right| + \left|(P_n - P_0) I(\bar{Q}_{b,n} > \eta)\right| \\ &\le \underbrace{\left|P_0 \left(I(\bar{Q}_{b,n} > \eta) - I(\bar{Q}_{b,0} > \eta)\right)\right|}_{\triangleq T_n(\eta)} + o_{P_0}(1), \end{aligned} \tag{3.12}$$

where we use the notation $Pf = E_P[f(O)]$ for any distribution P and function f. Let

$Z_n(\eta)(w) \triangleq \left(I(\bar{Q}_{b,n}(w) > \eta) - I(\bar{Q}_{b,0}(w) > \eta)\right)^2$. The following display holds for all $q > 0$:

$$\begin{aligned}
T_n(\eta) &\leq P_0 Z_n(\eta) \\
&= P_0 Z_n(\eta) I(|\bar{Q}_{b,0} - \eta| > q) + P_0 Z_n(\eta) I(|\bar{Q}_{b,0} - \eta| \leq q) \\
&= P_0 Z_n(\eta) I(|\bar{Q}_{b,0} - \eta| > q) + P_0 Z_n(\eta) I(0 < |\bar{Q}_{b,0} - \eta| \leq q) \qquad (3.13) \\
&\leq P_0 Z_n(\eta) I(|\bar{Q}_{b,n} - \bar{Q}_{b,0}| > q) + P_0 Z_n(\eta) I(0 < |\bar{Q}_{b,0} - \eta| \leq q) \qquad (3.14) \\
&\leq \mathrm{Pr}_0 \left(|\bar{Q}_{b,n} - \bar{Q}_{b,0}| > q\right) + \mathrm{Pr}_0 \left(0 < |\bar{Q}_{b,0} - \eta| \leq q\right) \\
&\leq \frac{P_0 |\bar{Q}_{b,n} - \bar{Q}_{b,0}|}{q} + \mathrm{Pr}_0 \left(0 < |\bar{Q}_{b,0} - \eta| \leq q\right).
\end{aligned}$$

Above (3.13) holds because C3) implies that $\mathrm{Pr}_0(\bar{Q}_{b,0} = \eta) = 0$, (3.14) holds because $Z_n(\eta) = 1$ implies that $|\bar{Q}_{b,n} - \bar{Q}_{b,0}| \geq |\bar{Q}_{b,0} - \eta|$, and the final inequality holds by Markov's inequality. The lemma assumes that $E_0|\bar{Q}_{b,n} - \bar{Q}_{b,0}| = o_{P_0}(1)$, and thus we can choose a sequence $q_n \downarrow 0$ such that

$$T_n(\eta) \leq \mathrm{Pr}_0 \left(0 < |\bar{Q}_{b,0} - \eta| \leq q_n\right) + o_{P_0}(1).$$

To see that the first term on the right is $o(1)$, note that $\mathrm{Pr}_0(\bar{Q}_{b,0} = \eta) = 0$ combined with the continuity of S_0 on N yield that, for n large enough,

$$\mathrm{Pr}_0 \left(0 < |\bar{Q}_{b,0} - \eta| \leq q_n\right) = S_0(-q_n + \eta) - S_0(q_n + \eta).$$

The right-hand side is $o(1)$, and thus $T_n(\eta) = o_{P_0}(1)$. Plugging this into (3.12) shows that $S_n(\eta) \to S_0(\eta)$ in probability. Recall that $\eta \in N$ was arbitrary.

Fix $\gamma > 0$. For γ small enough, $\eta_0 - \gamma$ and $\eta_0 + \gamma$ are contained in N. Thus $S_n(\eta_0 - \gamma) \to S_0(\eta_0 - \gamma)$ and $S_n(\eta_0 + \gamma) \to S_0(\eta_0 + \gamma)$ in probability. Further, $S_0(\eta_0 - \gamma) > \kappa$ by the definition of η_0 and $S_0(\eta_0 + \gamma) < \kappa$ by Condition C2). It follows that, with probability approaching 1, $S_n(\eta_0 - \gamma) > \kappa$ and $S_n(\eta_0 + \gamma) < \kappa$. But $|\eta_n - \eta_0| > \gamma$ implies that $S_n(\eta_0 - \gamma) \leq \kappa$ or $S_n(\eta_0 + \gamma) > \kappa$, and thus $|\eta_n - \eta_0| \leq \gamma$ with probability approaching 1. Thus $\eta_n \to \eta_0$ in probability, and $\tau_n \to \tau_0$ by the continuous mapping theorem. □

Chapter 4

Inference for Infinite-Dimensional Parameters

4.1 Introduction

In this chapter, we present a general confidence procedure for testing that a possibly unknown (but estimable function) has the same distribution as another possibly unknown (but estimable) function when applied to the observed data structure. To reduce the risk of deriving misleading conclusions due to model misspecification, it is appealing to employ flexible statistical learning tools to estimate these unknown functions. Unfortunately, inference is usually extremely difficult when such techniques are used, because the resulting estimators tend to be highly irregular. In the context of individualized medicine, this procedure can be used to test whether or not the conditional average treatment effect function, which is unknown but estimable from a sample of i.i.d. observations, is equal almost surely to zero. We will formally define the conditional average treatment effect, also known as the blip function, later in this Introduction. Throughout this chapter, we refer to this example as our Motivating Example. In fact, the methods of this chapter can be used to test if the blip function is almost surely equal to any fixed function f. Thus, as we note in the discussion, this allows the user to construct a confidence set for this infinite-dimensional parameter.

Because most of this chapter is written in a more general setting than previous chapters, we change a few of the notation restrictions introduced in Chapter 1. We still assume that the observed data is given by $O_1, \ldots, O_n$ drawn i.i.d. from some distribution P_0, but we now assume that the model $\mathcal{M}$ to which P_0 belongs is fully (locally) nonparametric. Furthermore, we remove the restriction that O must take the form (W, A, Y) for covariates W, binary treatment A, and outcome Y: while the support $\mathcal{O}$ of O takes this form in the examples we consider in this chapter, our general presentation of this approach does not make any assumptions on $\mathcal{O}$.

We now formulate the problem that this chapter considers statistically. Suppose that $P \mapsto S_P$ and $P \mapsto R_P$ are parameters mapping from $\mathcal{M}$ onto the space of univariate bounded real-valued measurable functions defined on $\mathcal{O}$. For brevity, we will write $R_0 \triangleq R_{P_0}$ and $S_0 \triangleq S_{P_0}$. Our objective is to test the null hypothesis

$$\mathcal{H}_0 : R_0(O) \stackrel{d}{=} S_0(O)$$

versus the complementary alternative $\mathcal{H}_1$: not $\mathcal{H}_0$, where O follows the distribution P_0 and the symbol $\stackrel{d}{=}$ denotes equality in distribution. We note that $R_0(O) \stackrel{d}{=} S_0(O)$ if $R_0 \equiv S_0$, i.e. $R_0(O) = S_0(O)$ almost surely, but not conversely. The case where $S_0 \equiv 0$ is of particular interest since then the null simplifies to $\mathcal{H}_0 : R_0 \equiv 0$. Because P_0 is unknown, R_0 and S_0 are not readily available. Nevertheless, the observed data can be used to estimate P_0 and hence each of R_0 and S_0. The approach we propose will apply to functionals within a specified class described later.

Before presenting our general approach, we describe some motivating examples. Consider the data structure $O = (W, A, Y)$, where W is a collection of covariates, A is binary treatment indicator, and Y is a bounded outcome, and suppose that O is distributed according to P.

Example 1 (Motivating Example): Testing a null conditional average treatment effect.

We have already briefly introduced this example, but now we do so with notation. Let R_P be the blip function, i.e.

$$R_P(o) \triangleq E_P(Y \mid A = 1, W = w) - E_P(Y \mid A = 0, W = w),$$

and let $S_P \equiv 0$. In this case the null hypothesis corresponds to the absence of a conditional average treatment effect.

Example 2: Testing for equality in distribution of regression functions in two populations.

Suppose the setting of the previous example, but where A represents membership to population 0 or 1. Let

$$R_P(o) \triangleq E_P(Y \mid A = 1, W = w),$$
$$S_P(o) \triangleq E_P(Y \mid A = 0, W = w).$$

In this case the null hypothesis corresponds to the outcome having the conditional mean functions, applied to a random draw of the covariate, having the same distribution in these two populations. We note here that our formulation considers selection of individuals from either population as random rather than fixed so that population-specific sample sizes (as opposed to the total sample size) are themselves random. The same interpretation could also be used for the previous example, now testing if the two regression functions are equivalent.

Example 3: Testing a null covariate effect on average response.

Suppose now that the data unit only consists of $O \triangleq (W, Y)$. If $R_P(o)$ is defined as $E_P(Y \mid W = w)$ and $S_P \equiv 0$, the null hypothesis corresponds to the outcome Y having conditional mean zero in all strata of covariates. This may be interesting when zero has a special importance for the outcome, such as when the outcome is the profit over some period.

Example 4: Testing a null variable importance.

Suppose again that $O \triangleq (W, Y)$ and $W \triangleq (W(1), W(2), \ldots, W(K))$. Denote by $W(-k)$ the vector $(W(i) : 1 \leq i \leq K, i \neq k)$. Setting $R_P(o) \triangleq E_P(Y \mid W = w)$ and $S_P(o) \triangleq E_P(Y \mid W(-k) = w(-k))$, the null hypothesis corresponds to $W(k)$ having null variable importance in the presence of $W(-k)$ with respect to the conditional mean of Y given W in the sense that $E_P(Y \mid W) = E_P(Y \mid W(-k))$ almost surely. This is true because if $R_0(W) \stackrel{d}{=} S_0(W(-k))$, the latter random variables have equal variance and so

$$\begin{aligned} E_0\{\mathrm{Var}_{P_0}[R_0(W) \mid W(-k)]\} &= \mathrm{Var}_{P_0}[R_0(W)] - \mathrm{Var}_{P_0}\{E_0[R_0(W) \mid W(-k)]\} \\ &= \mathrm{Var}_{P_0}[R_0(W)] - \mathrm{Var}_{P_0}[S_0(W(-k))] = 0 , \end{aligned}$$

implying that $\mathrm{Var}_{P_0}[R_0(W) \mid W(-k)] = 0$ almost surely. Thus, a test of $R_P(O)$ equal in distribution to $S_P(O)$ is equivalent to a test of almost sure equality between R_P and S_P in this example. We will show in Section 4.6 that our approach cannot be directly applied to this example, but that a simple extension yields a valid test.

Gretton et al. (2006) investigated the related problem of testing equality between two distributions in a two-sample problem. They proposed estimating the maximum mean discrepancy (hereafter referred to as MMD), a non-negative numeric summary that equals zero if and only if the two distributions are equal. In this chapter, we also utilize the MMD as a parsimonious summary of equality but consider the more general problem wherein the null hypothesis relies on unknown functions R_0 and S_0 indexed by the data-generating distribution P_0.

Other investigators have proposed omnibus tests of hypotheses of the form $\mathcal{H}_0$ versus $\mathcal{H}_1$ in the literature. In the setting of our Motivating Example, the work presented in Racine et al. (2006) and Lavergne et al. (2015) is particularly relevant. The null hypothesis of interest in these papers consists of the equality $E_0[Y|A, W] = E_0[Y|W]$ holding almost surely. If individuals have a nontrivial probability of receiving treated in all strata of covariates, this null hypothesis is equivalent to $\mathcal{H}_0$. In both of these papers, kernel smoothing is used to estimate the required regression functions. Therefore, key smoothness assumptions are needed for their methods to yield valid conclusions. The method we present does not hinge on any particular class of estimators and therefore does not rely on this condition.

The chapter is organized as follows. In Section 4.2, we formally present our parameter of interest, the squared MMD between two unknown functions. In Section 4.3 we establish asymptotic representations for this parameter based on its higher-order differentiability, which, as we formally establish, holds even when the MMD involves estimation of unknown nuisance parameters. In Section 4.4, we discuss estimation of this parameter, discuss the corresponding hypothesis test and study its asymptotic behavior under the null. We study the consistency of our proposed test under fixed and local alternatives in Section 4.5. We revisit our Motivating Example in Section 4.6. In Section 4.7, we present results from a simulation study to illustrate the finite-sample performance of our test, and we end with concluding remark in Section 4.8.

All proofs of our results can be found in Section 4.9.

4.2 Definition of the maximum mean discrepancy

For a distribution P and mappings T and U, we define

$$\Phi^{TU}(P) \triangleq \iint e^{-[T_P(o_1)-U_P(o_2)]^2} dP(o_1)dP(o_2) \tag{4.1}$$

and set $\Theta(P) \triangleq \Phi^{RR}(P) - 2\Phi^{RS}(P) + \Phi^{SS}(P)$. The MMD between the distributions of $R_P(O)$ and $S_P(O)$ when $O \sim P$ is given by $\sqrt{\Theta(P)}$ and is always well-defined because $\Theta(P)$ is non-negative. Indeed, denoting by θ_0 the true parameter value $\Theta(P_0)$, Theorem 3 of Gretton et al. (2006) establishes that θ_0 equals zero if $\mathcal{H}_0$ holds and is otherwise strictly positive. Though the study in Gretton et al. (2006) is restricted to two-sample problems, their proof of this result is only based upon properties of Θ and therefore holds regardless of the sample collected. Their proof relies on the fact that two random variables X and Y with compact support are equal in distribution if and only if $E[f(Y)] = E[f(X)]$ for every continuous function f, and uses techniques from the theory of Reproducing Kernel Hilbert Spaces (see, e.g., Berlinet and Thomas-Agnan, 2011, for a general exposition). We invite interested readers to consult Gretton et al. (2006) – and, in particular, Theorem 3 therein – for additional details. The definition of the MMD we utilize is based on the univariate Gaussian kernel with unit bandwidth, which is appropriate in view of Steinwart (2002). The results we present in this chapter can be generalized to the MMD based on a Gaussian kernel of arbitrary bandwidth by simply rescaling the mappings R and S.

4.3 Differentiability of the parameter of interest

Review of first- and second-order pathwise differentiability

We review the definitions of first- and second-order pathwise differentiability before showing that the MMD parameter Θ is sufficiently smooth to have these derivatives exist. Define the following fluctuation submodel:

$$dP_t(o) \triangleq \left(1 + th_1(o) + t^2h_2(o)\right) dP_0(o),$$
$$\text{where } P_0h_j = 0 \text{ and } \sup_{o\in\mathcal{O}} |h_j(o)| < \infty, j = 1, 2.$$

The function h_1 is a score, and the closure of the linear span of all such scores yields the unrestricted tangent space $L_0^2(P_0)$, i.e. the set of P_0 mean zero functions in $L^2(P_0)$.

Let $\theta_t \triangleq \Theta(P_t)$. The parameter Θ is called (first-order) pathwise differentiable at P_0 if there exists a $D_1^\Theta \in L_0^2(P_0)$ such that

$$\theta_t - \theta_0 = tP_0 D_1^\Theta h_1 + o(t).$$

We call D_1^Θ the first-order canonical gradient of Θ at P_0, where we note that $D_1^\Theta(O)$ is almost surely unique because $\mathcal{M}$ is nonparametric. The canonical gradient D_1^Θ depends on P_0 but this is omitted in the notation because we will only discuss pathwise differentiability at P_0.

A function $f : \mathcal{O}^2 \to \mathbb{R}$ is called (P) one-degenerate if it is symmetric and $Pf(o, \cdot) = 0$. We will use the notation $P^2 f = E_{P^2}[f(O_1, O_2)]$. The parameter Θ is called second-order pathwise differentiable at P_0 if there exists some symmetric, one-degenerate, P_0^2 square integrable function D_2^Θ such that

$$\theta_t - \theta_0 = tP_0 D_1^\Theta h_1 + \frac{1}{2}t^2 P_0 D_1^\Theta h_2 + \frac{1}{2}t^2 \int\int D_2^\Theta(o_1, o_2) h_1(o_1) h_1(o_2) dP_0(o_2) dP_0(o_1) + o(t^2).$$

First-order differentiability

To develop a test of $\mathcal{H}_0$, we will first construct an estimator θ_n of θ_0. In order to avoid restrictive model assumptions, we wish to use flexible estimation techniques in estimating P_0 and therefore θ_0. To control the operating characteristics of our test, it will be crucial to understand how to generate a parametric-rate estimator of θ_0. For this purpose, it is informative to first investigate the pathwise differentiability of Θ as a parameter from $\mathcal{M}$ to $\mathbb{R}$.

So far, we have not specified restrictions on the mappings R_P and S_P. However, in our developments, we will require these mappings to satisfy certain regularity conditions. Specifically, we will restrict our attention to elements of the class $\mathcal{S}$ of all mappings T for which there exists some measurable function X^T defined on $\mathcal{O}$ such that

(S1) T_P is a measurable mapping with domain $\{X^T(o) : o \in \mathcal{O}\}$ and range contained in $[-b, b]$ for some $0 \leq b < \infty$ independent of P;

(S2) there exists some $\delta > 0$ and a set $\mathcal{O}_1 \subseteq \mathcal{O}$ with $P_0(\mathcal{O}_1) = 1$ such that, for all $(o, t_1) \in \mathcal{O}_1 \times (-\delta, \delta)$, $t \mapsto T_{P_t}(x^T)$ is twice differentiable at t_1 with uniformly bounded first and second derivatives;

(S3) for any $P \in \mathcal{M}$ and submodel $dP_t/dP = 1 + th$ for uniformly bounded h with $Ph = 0$, there exists a function $D_P^T : \mathcal{O} \to \mathbb{R}$ uniformly bounded (in P and o) such that $\int D_P^T(o) dP(o|x^T) = 0$ for almost all $o \in \mathcal{O}$ and

$$\left.\frac{d}{dt} T_{P_t}(x^T)\right|_{t=0} = \int D_P^T(o) h(o) dP(o|x^T) \ .$$

Condition (S1) ensures that T is bounded and only relies on a summary measure of an observation O. Condition (S2) ensures that we will be able to interchange differentiation and integration when needed. Condition (S3) is a conditional (and weaker) version of pathwise differentiability in that the typical inner product representation only needs to hold for the conditional distribution of O given X^T under P_0. This concept is discussed further below. We will verify in Section 4.6 that these conditions hold in the context of the motivating examples presented earlier.

Remark 1. As a caution to the reader, we warn that simultaneously satisfying (S1) and (S3) may at times be restrictive. For example, if the observed data unit is $O \triangleq (W(1), W(2), Y)$, the parameter

$$T_P(o) \triangleq E_P\left[Y \mid W(1) = w(1), W(2) = w(2)\right] - E_P\left[Y \mid W(1) = w(1)\right]$$

cannot generally satisfy both conditions. Section 4.6 provides a means to tackle this problem using the techniques we have developed. In concluding remarks, we discuss a weakening of our conditions, notably by replacing $\mathcal{S}$ by the linear span of elements in $\mathcal{S}$. Consideration of this larger class significantly complicates the form of the estimator we propose in Section 4.4. □

We are now in a position to discuss the pathwise differentiability of Θ. For any elements $T, U \in \mathcal{S}$, we define

$$\Gamma_P^{TU}(o_1, o_2) \triangleq \Bigg[2\left[T_P(o_1) - U_P(o_2)\right]\left[D_P^U(o_2) - D_P^T(o_1)\right] + 1 \\ - \left\{4\left[T_P(o_1) - U_P(o_2)\right]^2 - 2\right\} D_P^T(o_1) D_P^U(o_2)\Bigg] e^{-[T_P(o_1) - U_P(o_2)]^2} .$$

and set $\Gamma_P \triangleq \Gamma_P^{RR} - \Gamma_P^{RS} - \Gamma_P^{SR} + \Gamma_P^{SS}$. Note that Γ_P is symmetric for any $P \in \mathcal{M}$. For brevity, we will write Γ_0^{TU} and Γ_0 to denote $\Gamma_{P_0}^{TU}$ and Γ_{P_0}, respectively. The following theorem characterizes the first-order behavior of Θ at an arbitrary $P \in \mathcal{M}$.

Theorem 11 (First-order pathwise differentiability of Θ over $\mathcal{M}$)**.** *If $R, S \in \mathcal{S}$, the parameter $\Theta : \mathcal{M} \to \mathbb{R}$ is pathwise differentiable at $P \in \mathcal{M}$ with first-order canonical gradient given by $D_1^\Theta(P)(o) \triangleq 2\left[\int \Gamma_P(o, o_2) dP(o_2) - \Theta(P)\right]$.*

Under some conditions, it is straightforward to construct an asymptotically linear estimator of θ_0 with influence function $D_1^\Theta(P_0)$, that is, an estimator θ_n of θ_0 such that

$$\theta_n - \theta_0 = \frac{1}{n}\sum_{i=1}^n D_1^\Theta(P_0)(O_i) + o_{P_0}(n^{-1/2}) .$$

For example, the one-step Newton-Raphson bias correction procedure (see, e.g., Pfanzagl, 1982) or targeted minimum loss-based estimation (see, e.g., van der Laan and Rose, 2011) can be used for this purpose. If the above representation holds and the variance of $D_1^{\Theta}(P_0)(O)$ is positive, then $\sqrt{n}\,(\theta_n - \theta_0) \rightsquigarrow N(0, \sigma_0^2)$, where the symbol $\rightsquigarrow$ denotes convergence in distribution and we write $\sigma_0^2 \triangleq P_0\left[D_1^{\Theta}(P_0)^2\right]$. If σ_0 is strictly positive and can be consistently estimated, Wald-type confidence intervals for θ_0 with appropriate asymptotic coverage can be constructed.

The situation is more challenging if $\sigma_0 = 0$. In this case, $\sqrt{n}\,(\theta_n - \theta_0) \to 0$ in probability and typical Wald-type confidence intervals will not be appropriate. Because $D_1^{\Theta}(P_0)(O)$ has mean zero under P_0, this happens if and only if $D_1^{\Theta}(P_0) \equiv 0$. The following lemma provides necessary and sufficient conditions under which $\sigma_0 = 0$.

Corollary 2 (First-order degeneracy under $\mathcal{H}_0$)**.** *If $R, S \in \mathcal{S}$, it will be the case that $\sigma_0 = 0$ if and only if either (i) $\mathcal{H}_0$ holds, or (ii) $R_0(O)$ and $S_0(O)$ are degenerate with $D_0^R \equiv D_0^S$.*

The above results rely in part on knowledge of D_0^R and D_0^S. It is useful to note that, in some situations, the computation of $D_P^T(o)$ for a given $T \in \mathcal{S}$ and $P \in \mathcal{M}$ can be streamlined. This is the case, for example, if $P \mapsto T_P$ is invariant to fluctuations of the marginal distribution of X^T, as it seems (S3) may suggest. Consider obtaining i.i.d. samples of increasing size from the conditional distribution of O given $X^T = x^T$ under P, so that all individuals have observed $X^T = x^T$. Consider the fluctuation submodel $dP_t(o|x^T) \triangleq [1 + th(o)]\, dP(o|x^T)$ for the conditional distribution, where h is uniformly bounded and $\int h(o)dP(o|x^T) = 0$. Suppose that (i) $P \mapsto T_P(x^T)$ is differentiable at $t = 0$ with respect to the above submodel and (ii) this derivative satisfies the inner product representation

$$\left.\frac{d}{dt}T_{P_t}(x^T)\right|_{t=0} = \int \tilde{D}_P^T(o|x^T)h(o)dP(o|x^T)$$

for some uniformly bounded function $o \mapsto \tilde{D}_P^T(o|x^T)$ with $\int \tilde{D}_P^T(o|x^T)dP(o|x^T) = 0$. If the above holds for all x^T, we may take $D_P^T(o) = \tilde{D}_P^T(o|x^T)$ for all o with $X^T(o) = x^T$. If D_P^T is uniformly bounded in P, (S3) then holds.

In summary, the above discussion suggests that, if T is invariant to fluctuations of the marginal distribution of X^T, (S3) can be expected to hold if there exists a regular, asymptotically linear estimator of each $T_P(x^T)$ under i.i.d. sampling from the conditional distribution of O given $X^T = x^T$ implied by P.

Remark 2. If T is invariant to fluctuations of the marginal distribution of X^T, one can also expect (S3) to hold if $P \mapsto \int T_P(X^T(o))dP(o)$ is pathwise differentiable with canonical gradient uniformly bounded in P and o in the model in which the marginal distribution of X is known. The canonical gradient in this model is equal to D_P^T. □

Second-order differentiability and asymptotic representation

As indicated above, if $\sigma_0 = 0$, the behavior of Θ around P_0 cannot be adequately characterized by a first-order analysis. For this reason, we must investigate whether Θ is second-order differentiable. As we discuss below, under $\mathcal{H}_0$, Θ is indeed second-order pathwise differentiable at P_0 and admits a useful second-order asymptotic representation.

Theorem 12 (Second-order pathwise differentiability under $\mathcal{H}_0$)**.** *If $R, S \in \mathcal{S}$ and $\mathcal{H}_0$ holds, the parameter $\Theta : \mathcal{M} \to \mathbb{R}$ is second-order pathwise differentiable at P_0 with second-order canonical gradient $D_2^\Theta(P_0) \triangleq 2\Gamma_0$.*

It is easy to confirm that Γ_0, and thus D_2^Θ, is one-degenerate under $\mathcal{H}_0$ in the sense that $\int \Gamma_0(o, o_2)dP_0(o_2) = \int \Gamma_0(o_1, o)dP_0(o_1) = 0$ for all o. This is shown as follows. For any $T, U \in \mathcal{S}$, the law of total expectation conditional on X^U and fact that $\int D_0^U(o)dP_0(o|x^U) = 0$ yields that

$$\int \Gamma_0^{TU}(o, o_2)dP_0(o_2) = \int \left\{1 - 2\left[T_0(o) - U_0(o_2)\right] D_0^T(o)\right\} e^{-[T_0(o)-U_0(o_2)]^2} dP_0(o_2) \ ,$$

where we have written Γ_0^{TU} to denote $\Gamma_{P_0}^{TU}$. Since $\int f(R_0(o))dP_0(o) = \int f(S_0(o))dP_0(o)$ for each measurable function f when $S_0(O) \stackrel{d}{=} T_0(O)$, this then implies that $\int \Gamma_0^{RS}(o, o_2)dP_0(o_2) = \int \Gamma_0^{RR}(o, o_2)dP(o_2)$ and $\int \Gamma_0^{SR}(o, o_2)dP_0(o_2) = \int \Gamma_0^{SS}(o, o_2)dP_0(o_2)$ under $\mathcal{H}_0$. Hence, it follows that $\int \Gamma_0(o, o_2)dP_0(o_2) = 0$ under $\mathcal{H}_0$ for any o.

If second-order pathwise differentiability held in a sufficiently uniform sense over $\mathcal{M}$, we would expect

$$\text{Rem}_P^\Theta \ \triangleq \ \Theta(P) - \Theta(P_0) - (P - P_0)D_1^\Theta(P) + \frac{1}{2}(P - P_0)^2 D_2^\Theta(P) \tag{4.2}$$

to be a third-order remainder term. However, second-order pathwise differentiable has only been established under the null, and in fact, it appears that Θ may not generally be second-order pathwise differentiable under the alternative. As such, D_2^Θ may not even be defined under the alternative. In writing (4.2), we either naively set $D_2^\Theta(P) \triangleq 2\Gamma_P$, which is not appropriately centered to be a candidate second-order gradient, or instead take D_2^Θ to be the centered extension $(o_1, o_2) \mapsto 2\left[\Gamma_P(o_1, o_2) - \int \Gamma_P(o_1, o)dP(o) - \int \Gamma_P(o, o_2)dP(o) + P^2\Gamma_P\right]$. Both of these choices yield the same expression above because the product measure $(P - P_0)^2$ is self-centering. The need for an extension renders it a priori unclear whether as P tends to P_0 the behavior of Rem_P^Θ is similar to what is expected under more global second-order pathwise differentiability. Using the fact that $\Theta(P) = P^2\Gamma_P$, we can simplify the expression in (4.2) to

$$\text{Rem}_P^\Theta = P_0^2\Gamma_P - \theta_0 \ . \tag{4.3}$$

As we discuss below, this remainder term can be bounded in a useful manner, which allows us to determine that it is indeed third-order.

For all $T \in \mathcal{S}$, $P \in \mathcal{M}$ and $o \in \mathcal{O}$, we define

$$\text{Rem}_P^T(o) \ \triangleq \ T_P(o) - T_0(o) + \int D_P^T(o_1)\left[dP(o_1|x^T) - dP_0(o_1|x^T)\right]$$

as the remainder from the linearization of T based on the conditional gradient D_P^T. Typically, $\text{Rem}_P^T(o)$ is a second-order term. Further consideration of this term in the context of our motivating examples is described in Section 4.6. Furthermore, we define

$$\begin{aligned} L_P^{RS}(o) &\triangleq \max\left\{|\operatorname{Rem}_P^R(o)|, |\operatorname{Rem}_P^S(o)|\right\} \\ M_P^{RS}(o) &\triangleq \max\left\{|R_P(o) - R_0(o)|, |S_P(o) - S_0(o)|\right\}. \end{aligned}$$

For any given function $f : \mathcal{O} \to \mathbb{R}$, we denote by $\|f\|_{p,P_0} \triangleq \left[\int |f(o)|^p dP_0(o)\right]^{1/p}$ the $L^p(P_0)$-norm and use the symbol $\lesssim$ to denote 'less than or equal to up to a positive multiplicative constant'. The following theorem provides an upper bound for the remainder term of interest.

Theorem 13 (Upper bounds on remainder term). *For each $P \in \mathcal{M}$, the remainder term admits the following upper bounds:*

$$\begin{aligned} \textit{Under } \mathcal{H}_0: \ |\operatorname{Rem}_P^\Theta| &\lesssim K_{0P} \triangleq \left\|L_P^{RS}\right\|_{2,P_0} \left\|M_P^{RS}\right\|_{2,P_0} + \left\|L_P^{RS}\right\|_{1,P_0}^2 + \left\|M_P^{RS}\right\|_{4,P_0}^4 \\ \textit{Under } \mathcal{H}_1: \ |\operatorname{Rem}_P^\Theta| &\lesssim K_{1P} \triangleq \left\|L_P^{RS}\right\|_{1,P_0} + \left\|M_P^{RS}\right\|_{2,P_0}^2 \ . \end{aligned}$$

To develop a test procedure, we will require an estimator of P_0, which will play the role of P in the above expressions. It is helpful to think of parametric model theory when interpreting the above result, with the understanding that certain smoothing methods, such as higher-order kernel smoothing, can achieve near-parametric rates in certain settings. In a parametric model, we could often expect $\left\|L_P^{RS}\right\|_{p,P_0}$ and $\left\|M_P^{RS}\right\|_{p,P_0}$ to be $O_{P_0}(n^{-1})$ and $O_{P_0}(n^{-1/2})$, respectively, for $p \geq 1$. Thus, the above theorem suggests that the approximation error may be $O_P(n^{-3/2})$ in a parametric model under $\mathcal{H}_0$. In some examples, it is reasonable to expect that $L_P^{RS} \equiv 0$ for a large class of distributions P. In such cases, the upper bound on Rem_P^Θ simplifies to $\left\|M_P^{RS}\right\|_{4,P_0}^4$ under $\mathcal{H}_0$, which under a parametric model is often $O_{P_0}(n^{-2})$.

4.4 Proposed test: formulation and inference under the null

Formulation of test

We begin by constructing an estimator of θ_0 from which a test can then be devised. Using the fact that $\Theta(P) = P^2\Gamma_P$, as implied by (4.3), we note that if Γ_0 were known, the U-statistic $\mathbb{U}_n\Gamma_0$ would be a natural estimator of θ_0, where $\mathbb{U}_n$ denotes the empirical measure that places equal probability mass on each of the $n(n-1)$ points (O_i, O_j) with $i \neq j$. In practice, Γ_0 is unknown and must be estimated. This leads to the estimator $\theta_n \triangleq \mathbb{U}_n\Gamma_n$, where we write $\Gamma_n \triangleq \Gamma_{\hat{P}_n}$ for some estimator $\hat{P}_n$ of P_0 based on the available data. Since a large value of θ_n is inconsistent with $\mathcal{H}_0$, we will reject $\mathcal{H}_0$ if and only if $\theta_n > c_n$ for some appropriately chosen cutoff c_n.

In the nonparametric model considered, it may be necessary, or at the very least desirable, to utilize a data-adaptive estimator $\hat{P}_n$ of P_0 when constructing Γ_n. Studying the large-sample properties of θ_n may then seem particularly daunting since at first glance we may be led to believe that the behavior of $\theta_n - \theta_0$ is dominated by $P_0^2\left(\Gamma_n - \Gamma_0\right)$. However, this is not the case. As we will see, under some conditions, $\theta_n - \theta_0$ will approximately behave like $(\mathbb{U}_n - P_0^2)\,\Gamma_0$. Thus, there will be no contribution of $\hat{P}_n$ to the asymptotic behavior of $\theta_n - \theta_0$. Though this result may seem counterintuitive, it arises because $\Theta(P)$ can be expressed as $P^2\Gamma_P$ with Γ_P a second-order gradient (or rather an extension thereof) up to a proportionality constant. More concretely, this surprising finding is a direct consequence of (4.3).

Remark 3. As further support that θ_n may indeed be expected to have good properties, even when a data-adaptive estimator $\hat{P}_n$ of P_0 has been used, we note that θ_n could also have been derived using a second-order one-step Newton-Raphson construction, as described in Robins et al. (2008). The latter is given by

$$\theta_{n,NR} \triangleq \Theta(\hat{P}_n) + P_n D_1^{\Theta}(\hat{P}_n) + \frac{1}{2}\mathbb{U}_n D_2^{\Theta}(\hat{P}_n)\ ,$$

where we use the centered extension of D_2^{Θ} as discussed in Section 4.3. Here and throughout, P_n denotes the empirical distribution. It is straightforward to verify that indeed $\theta_n = \theta_{n,NR}$. □

Inference under the null

Asymptotic behavior

For each $P \in \mathcal{M}$, we let $\tilde{\Gamma}_P$ be the P_0-centered modification of Γ_P given by

$$(o_1, o_2) \mapsto \tilde{\Gamma}_P(o_1, o_2) \triangleq \Gamma_P(o_1, o_2) - \int \Gamma_P(o_1, o) dP_0(o) - \int \Gamma_P(o, o_2) dP_0(o) + P_0^2\Gamma_P$$

and denote $\tilde{\Gamma}_{P_0}$ by $\tilde{\Gamma}_0$. While $\tilde{\Gamma}_0 = \Gamma_0$ under $\mathcal{H}_0$, this is not true more generally. Below, we use Rem_n^{Θ} and $\tilde{\Gamma}_n$ to respectively denote Rem_P^{Θ} and $\tilde{\Gamma}_P$ evaluated at $P = \hat{P}_n$. Straightforward algebraic manipulations allows us to write

$$\begin{aligned} \theta_n - \theta_0 &= \mathbb{U}_n\Gamma_n - \theta_0 \;=\; \mathbb{U}_n\Gamma_n - P_0^2\Gamma_n + P_0^2\Gamma_n - \theta_0 \\ &= \left(\mathbb{U}_n - P_0^2\right)\Gamma_n + \mathrm{Rem}_n^{\Theta} \\ &= \mathbb{U}_n\Gamma_0 + 2\left(P_n - P_0\right)P_0\Gamma_n + \mathbb{U}_n\left(\tilde{\Gamma}_n - \Gamma_0\right) + \mathrm{Rem}_n^{\Theta}\ . \end{aligned} \tag{4.4}$$

Our objective is to show that $n\left(\theta_n - \theta_0\right)$ behaves like $n\mathbb{U}_n\Gamma_0$ as n gets large under $\mathcal{H}_0$. In view of (4.4), this will be true, for example, under conditions ensuring that

C1) $n(P_n - P_0)P_0\Gamma_n = o_{P_0}(1)$ (empirical process and consistency conditions);

C2) $n\mathbb{U}_n\left(\tilde{\Gamma}_n - \Gamma_0\right) = o_{P_0}(1)$ (U-process and consistency conditions);

C3) $n\,\mathrm{Rem}_n^{\Theta} = o_{P_0}(1)$ (consistency and rate conditions).

We have already argued that C3) is reasonable in many examples of interest, including those presented in this chapter. Nolan and Pollard (1987, 1988) developed a formal theory that controls terms of the type appearing in C2). In the Supplementary Material of Luedtke et al. (2015), we restate specific results from these authors which are useful to to study C2). Finally, the following lemma gives sufficient conditions under which C1) holds. We first set $K_{1n} \triangleq \left\|L_{\hat{P}_n}^{RS}\right\|_{1,P_0} + \left\|M_{\hat{P}_n}^{RS}\right\|_{2,P_0}^2$.

Lemma 5 (Sufficient conditions for C1)). *Suppose that* $o_1 \mapsto \int \Gamma_n(o_1, o)dP_0(o)/K_{1n}$*, defined to be zero if* $K_{1n} = 0$*, belongs to a* P_0*-Donsker class with probability tending to* 1*. Then, under* $\mathcal{H}_0$*,*

$$(P_n - P_0)P_0\Gamma_n = O_{P_0}\left(\frac{K_{1n}}{\sqrt{n}}\right)$$

and thus C1) holds whenever $K_{1n} = o_{P_0}(n^{-1/2})$*.*

The following theorem describes the asymptotic distribution of $n\theta_n$ under the null hypothesis whenever conditions C1), C2) and C3) are satisfied.

Theorem 14 (Asymptotic distribution under $\mathcal{H}_0$). *Suppose that C1), C2) and C3) hold. Then, under* $\mathcal{H}_0$*,*

$$n\theta_n = n\mathbb{U}_n\Gamma_0 + o_{P_0}(1) \rightsquigarrow \sum_{k=1}^{\infty} \lambda_k\left(Z_k^2 - 1\right),$$

where $\{\lambda_k\}_{k=1}^{\infty}$ *are the eigenvalues of the integral operator* $h(o) \mapsto \int \Gamma_0(o_1, o)h(o)dP_0(o_1)$ *repeated according to their multiplicity, and* $\{Z_k\}_{k=1}^{\infty}$ *is a sequence of independent standard normal random variables. Furthermore, all of these eigenvalues are nonnegative under* $\mathcal{H}_0$*.*

We note that by employing a sample splitting procedure – namely, estimating Γ_0 on one portion of the sample and constructing the U-statistic based on the remainder of the sample – it is possible to eliminate the U-process conditions required for C2). In such a case, satisfaction of C2) only requires convergence of $\tilde{\Gamma}_n$ to Γ_0 with respect to the $L^2(P_0^2)$-norm.

Estimation of the test cutoff

As indicated above, our test consists of rejecting $\mathcal{H}_0$ if and only if θ_n is larger than some cutoff c_n. We wish to select c_n to yield a non-conservative test at level $\alpha \in (0, 1)$. In view of Theorem 14, denoting by $q_{1-\alpha}$ the $1-\alpha$ quantile of the described limit distribution, the cutoff c_n should be chosen to be $q_{1-\alpha}/n$. We thus reject $\mathcal{H}_0$ if and only if $n\theta_n > q_{1-\alpha}$. As described in the following corollary, $q_{1-\alpha}$ admits a very simple form when $S_P \equiv 0$ for all P.

Corollary 3 (Asymptotic distribution under $\mathcal{H}_0$, S degenerate)**.** *Suppose that C1), C2) and C3) hold, that $S_P \equiv 0$ for all $P \in \mathcal{M}$, and that $\sigma_R^2 \triangleq \mathrm{Var}_{P_0}\left[D_0^R(O)\right] > 0$. Then, under $\mathcal{H}_0$,*

$$\frac{n\theta_n}{2\sigma_R^2} \rightsquigarrow Z^2 - 1,$$

where Z is a standard normal random variable. It follows then that $q_{1-\alpha} = 2\sigma_R^2(z_{1-\alpha/2}^2 - 1)$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

The above corollary gives an expression for $q_{1-\alpha}$ that can easily be consistently estimated from the data. In particular, one can use $\hat{q}_{1-\alpha} \triangleq 2(z_{1-\alpha/2}^2 - 1)P_n D^R(\hat{P}_n)^2$ as an estimator of $q_{1-\alpha}$, whose consistency can be established under a Glivenko-Cantelli and consistency condition on the estimator of D_0^R. However, in general, such a simple expression will not exist. Gretton et al. (2009) proposed estimating the eigenvalues ν_k of the centered Gram matrix and then computing $\hat{\lambda}_k \triangleq \nu_k/n$. In our context, the eigenvalues ν_k are those of the $n \times n$ matrix $G \triangleq \{G_{ij}\}_{1\le i,j\le n}$ with entries defined as

$$G_{ij} \;\triangleq\; \Gamma_n(O_i,O_j) - \frac{1}{n}\sum_{k=1}^{n}\Gamma_n(O_k,O_j) - \frac{1}{n}\sum_{\ell=1}^{n}\Gamma_n(O_i,O_\ell) + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{\ell=1}^{n}\Gamma_n(O_k,O_\ell)\ . \quad (4.5)$$

Given these n eigenvalue estimates $\hat{\lambda}_1, ..., \hat{\lambda}_n$, one could then simulate from $\sum_{k=1}^n \hat{\lambda}_k(Z_k^2 - 1)$ to approximate $\sum_{k=1}^\infty \lambda_k(Z_k^2 - 1)$. While this seems to be a plausible approach, a formal study establishing regularity conditions under which this procedure is valid is beyond the scope of this chapter. We note that it also does not fall within the scope of results in Gretton et al. (2009) since their kernel does not depend on estimated nuisance parameters. We refer the reader to Franz (2006) for possible sufficient conditions under which this approach may be valid.

In practice, it suffices to give a data-dependent asymptotic upper bound on $q_{1-\alpha}$. We will refer to $\hat{q}_{1-\alpha}^{ub}$, which depends on P_n, as an asymptotic upper bound of $q_{1-\alpha}$ if

$$\limsup_{n\to\infty} \mathrm{Pr}_0\left(n\theta_n > \hat{q}_{1-\alpha}^{ub}\right) \;\le\; 1-\alpha\ . \quad (4.6)$$

If $q_{1-\alpha}$ is consistently estimated, one possible choice of $\hat{q}_{1-\alpha}^{ub}$ is this estimate of $q_{1-\alpha}$ – the inequality above would also become an equality provided the conclusion of Theorem 14 holds. It is easy to derive a data-dependent upper bound with this property using Chebyshev's inequality. To do so, we first note that

$$\mathrm{Var}_{P_0}\left[\sum_{k=1}^\infty \lambda_k\left(Z_k^2 - 1\right)\right] \;=\; \sum_{k=1}^\infty \lambda_k^2\,\mathrm{Var}_{P_0}\left(Z_k^2\right) \;=\; 2\sum_{k=1}^\infty \lambda_k^2 \;=\; 2P_0^2\Gamma_0^2\ ,$$

where we have interchanged the variance operation and the limit using the L^2 martingale convergence theorem and the last equality holds because λ_k, $k = 1, 2, \ldots$, are the eigenvalues of the Hilbert-Schmidt integral operator with kernel $\tilde{\Gamma}_0$. Under mild regularity conditions, $P_0^2\Gamma_0^2$ can be consistently estimated using $\mathbb{U}_n\Gamma_n^2$. Provided $P_0^2\Gamma_0^2 > 0$, we find that

$$\left(2\mathbb{U}_n\Gamma_n^2\right)^{-1/2} n\theta_n \rightsquigarrow \left(2P_0^2\Gamma_0^2\right)^{-1/2} \sum_{k=1}^{\infty} \lambda_k \left(Z_k^2 - 1\right) \ , \tag{4.7}$$

where the limit variate has mean zero and unit variance. The following theorem gives a valid choice of $\hat{q}_{0.95}^{ub}$.

Theorem 15. *Suppose that C1), C2) and C3) hold. Then, under* $\mathcal{H}_0$ *and provided* $\mathbb{U}_n\Gamma_n^2 \to P_0^2\Gamma_0^2 > 0$ *in probability,* $\hat{q}_{0.95}^{ub} \triangleq 6.2 \cdot (\mathbb{U}_n\Gamma_n^2)^{1/2} > q_{0.95}$ *is a valid upper bound in the sense of (4.6).*

The proof of the result follows immediately by noting that $P(X > t) \leq (1+t^2)^{-1}$ for any random variable X with mean zero and unit variance in view of the one-sided Chebyshev's inequality. This illustrates concretely that we can obtain a consistent test that controls type I error. In practice, we recommend either using the result of Corollary 3 whenever possible or estimating the eigenvalues of the matrix in (4.5). Nonetheless, we generally recommend either using the result of Corollary 3 whenever possible or estimating the eigenvalues of the matrix in (4.5).

We note that the condition $\sigma_R^2 > 0$ holds in many but not all examples of interest. Fortunately, the plausibility of this assumption can be evaluated analytically. In Section 4.6, we show that this condition does not hold in Example 4 and provide a way forward despite this.

4.5 Asymptotic behavior under the alternative

Consistency under a fixed alternative

We present two analyses of the asymptotic behavior of our test under a fixed alternative. The first relies on $\hat{P}_n$ providing a good estimate of P_0. Under this condition, we give an interpretable limit distribution that provides insight into the behavior of our estimator under the alternative. As we show, surprisingly, $\hat{P}_n$ need not be close to P_0 to obtain an asymptotically consistent test, even if the resulting estimate of θ_0 is nowhere near the truth. In the second analysis, we give more general conditions under which our test will be consistent under $\mathcal{H}_1$.

Nuisance functions have been estimated well

As we now establish, our test has power against all alternatives P_0 except for the fringe cases discussed in Corollary 2 with Γ_0 one-degenerate. We first note that

$$\theta_n - \theta_0 \ = \ \mathbb{U}_n\Gamma_n - \theta_0 \ = \ 2(P_n - P_0)P_0\Gamma_n + \mathbb{U}_n\tilde{\Gamma}_n + \mathrm{Rem}_P^{\Theta} \ .$$

When scaled by $\sqrt{n}$, the leading term on the right-hand side follows a mean zero normal distribution under regularity conditions. The second summand is typically $O_{P_0}(n^{-1})$ under certain conditions, for example, on the entropy of the class of plausible realizations of the random function $(o_1, o_2) \mapsto \Gamma_n(o_1, o_2)$ (Nolan and Pollard, 1987, 1988). In view of the second statement in Theorem 13, the third summand is a second-order term that will often be negligible, even after scaling by $\sqrt{n}$. As such, under certain regularity conditions, the leading term in the representation above determines the asymptotic behavior of θ_n, as described in the following theorem.

Theorem 16 (Asymptotic distribution under $\mathcal{H}_1$)**.** *Suppose that $K_{1n} = o_{P_0}(n^{-1/2})$, that $\mathbb{U}_n\tilde{\Gamma}_n = o_{P_0}(n^{-1/2})$, and furthermore, that $o \mapsto \int \Gamma_n(o_1, o)dP_0(o)$ belongs to a fixed P_0-Donsker class with probability tending to 1 while $\|P_0\left(\Gamma_n - \Gamma_0\right)\|_{2,P_0} = o_{P_0}(1)$. Under $\mathcal{H}_1$, we have that $\sqrt{n}\left(\theta_n - \theta_0\right) \rightsquigarrow N\left(0, \tau^2\right)$, where $\tau^2 \triangleq 4\,\mathrm{Var}_{P_0}\left[\int \Gamma_0(O, o)dP_0(o)\right]$.*

In view of the results of Section 4.2, τ^2 coincides with σ_0^2, the efficiency bound for regular, asymptotically linear estimators in a nonparametric model. Hence, θ_n is an asymptotically efficient estimator of θ_0 under $\mathcal{H}_1$.

The following corollary is trivial in light of Theorem 16. It establishes that the test $n\theta_n > \hat{q}^{ub}_{1-\alpha}$ is consistent against (essentially) all alternatives provided the needed components of the likelihood are estimated sufficiently well.

Corollary 4 (Consistency under a fixed alternative)**.** *Suppose the conditions of Theorem 16. Furthermore, suppose that $\tau^2 > 0$ and $\hat{q}^{ub}_{1-\alpha} = o_{P_0}(n)$. Then, under $\mathcal{H}_1$, the test $n\theta_n > \hat{q}^{ub}_{1-\alpha}$ is consistent in the sense that*

$$\lim_{n\to\infty} \mathrm{Pr}_0\left(n\theta_n > \hat{q}^{ub}_{1-\alpha}\right) = 1 \ .$$

The requirement that $\hat{q}^{ub}_{1-\alpha} = o_{P_0}(n)$ is very mild given that $q_{1-\alpha}$ will be finite whenever $R, S \in \mathcal{S}$. As such, we would not expect $\hat{q}^{ub}_{1-\alpha}$ to get arbitrarily large as sample size grows, at least beyond the extent allowed by our corollary. This suggests that most non-trivial upper bounds satisfying (4.6) will yield a consistent test.

Nuisance functions have not been estimated well

We now consider the case where the nuisance functions are not estimated well, in the sense that the consistency conditions of Theorem 16 do not hold. In particular, we argue that failure of these conditions does not necessarily undermine the consistency of our test. Let $\hat{q}^{ub}_{1-\alpha}$ be the estimated cutoff for our test, and suppose that $\hat{q}^{ub}_{1-\alpha} = o_{P_0}(n)$. Suppose also that $P_0^2\Gamma_n$ is asymptotically bounded away from zero in the sense that, for some $\delta > 0$, $\mathrm{Pr}_0\left(P_0^2\Gamma_n > \delta\right)$ tends to one. This condition is reasonable given that $P_0^2\Gamma_0 > 0$ under $\mathcal{H}_1$ and $\hat{P}_n$ is nevertheless a (possible inconsistent) estimator of P_0. Assuming that $\mathbb{U}_n\Gamma_n = O_{P_0}(n^{-1/2})$, which is true under entropy conditions on Γ_n (Nolan and Pollard, 1987, 1988), we have that

$$\Pr_0\left(n\theta_n > \hat{q}^{ub}_{1-\alpha}\right) \;=\; \Pr_0\left(\sqrt{n}\mathbb{U}_n\Gamma_n > \frac{\hat{q}^{ub}_{1-\alpha}}{\sqrt{n}} - \sqrt{n}P_0^2\Gamma_n\right) \longrightarrow 1\ .$$

We have accounted for the random $n^{-1/2}\hat{q}^{ub}_{1-\alpha}$ term as in the proof of Corollary 4. Of course, this result is less satisfying than Theorem 16, which provides a concrete limit distribution.

Consistency under a local alternative

We consider local alternatives of the form

$$dQ_n(o) = \left[1 + n^{-1/2}h_n(o)\right] dP_0(o),$$

where $h_n \to h$ in $L^2_0(P_0)$ for some non-degenerate h and P_0 satisfies the null hypothesis $\mathcal{H}_0$. Suppose that the conditions of Theorem 14 hold. By Theorem 2.1 of Gregory (1977), we have that

$$n\mathbb{U}_n\Gamma_0 \overset{Q_n}{\rightsquigarrow} \sum_{k=1}^{\infty}\lambda_k\left[\left(Z_k + \langle f_k, h\rangle\right)^2 - 1\right],$$

where $\mathbb{U}_n$ is the U-statistic empirical measure from a sample of size n drawn from Q_n, $\langle\cdot,\cdot\rangle$ is the inner product in $L^2(P_0)$, Z_k and λ_k are as in Theorem 14, and f_k is the eigenfunction corresponding to eigenvalue λ_k described in Theorem 14. By the contiguity of Q_n, the conditions of Theorem 14 yield that the result above also holds with $\mathbb{U}_n\Gamma_0$ replaced by $\mathbb{U}_n\Gamma_n$, our estimator applied to a sample of size n drawn from Q_n.

If each λ_k is non-negative, the limiting distribution under Q_n stochastically dominates the asymptotic distribution under P_0, and furthermore, if $\langle f_k, h\rangle \neq 0$ for some k with $\lambda_k > 0$, this dominance is strict. It is straightforward to show that, under the conditions of Theorem 14, the above holds if and only if $\liminf_n \sqrt{n}\Theta(Q_n) > 0$, that is, if the sequence of alternatives is not too hard. Suppose that $\hat{q}_{1-\alpha}$ is a consistent estimate of $q_{1-\alpha}$. By Le Cam's third lemma, $\hat{q}_{1-\alpha}$ is consistent for $q_{1-\alpha}$ even when the estimator is computed on samples of size n drawn from Q_n rather than P_0. This proves the following theorem.

Theorem 17 (Consistency under a local alternative)**.** *Suppose that the conditions of Theorem 14 hold. Then, under $\mathcal{H}_0$ and provided $\liminf_{n\to\infty}\sqrt{n}\Theta(Q_n) > 0$, the proposed test is locally consistent in the sense that $\lim_{n\to\infty} Q_n\left(n\theta_n > \hat{q}_{1-\alpha}\right) > \alpha$.*

4.6 Illustrations

We now return to each of our examples. We first show that Examples 1, 2 and 3 satisfy the regularity conditions described in Section 4.2. Specifically, we show that all involved parameters R and S belong to $\mathcal{S}$ under reasonable conditions. Furthermore, we determine explicit remainder terms for the asymptotic representation used in each example and describe conditions under which these remainder terms are negligible. For any $T \in \mathcal{S}$, we will use the shorthand notation $\dot{T}_{\tilde{t}}(x^T) \triangleq \frac{d}{dt} T_{P_t}(x^T)\big|_{t=\tilde{t}}$ for $\tilde{t}$ in a neighborhood of zero.

Example 1 (Motivating Example) (Continued).

The parameter S with $S_P \equiv 0$ belongs to $\mathcal{S}$ trivially, with $D_P^S \equiv 0$. Condition (S1) holds with $x^R(o) = w$. Condition (S2) holds using that

$$R_t(w) = \sum_{a=0}^{1} (-1)^{a+1} \int y \left\{ \frac{1 + t h_1(w,a,y) + t^2 h_2(w,a,y)}{1 + t E_0[h_1(w,A,Y)] + t^2 E_0[h_2(w,A,Y)]} \right\} dP_0(y|a,w) \ . \quad (4.8)$$

Since we must only consider h_1 and h_2 uniformly bounded, for t sufficiently small, we see that $R_t(w)$ is twice continuously differentiable with uniformly bounded derivatives. Condition (S3) is satisfied by

$$D_P^R(o) \triangleq \frac{2a-1}{P\left(A = a \mid W = w\right)} \left\{ y - E_P\left[Y \mid A = a, W = w\right] \right\}$$

and $D_P^S \equiv 0$. If $\min_a P\left(A = a \mid W\right)$ is bounded away from zero with probability 1 uniformly in P, it follows that $(P, o) \mapsto D_0^R(o)$ is uniformly bounded.

Clearly, we have that $\mathrm{Rem}_P^S \equiv 0$. We can also verify that $\mathrm{Rem}_P^R(o)$ equals

$$\sum_{\tilde{a}=0}^{1} (-1)^{\tilde{a}} E_0 \left\{ \left[1 - \frac{P_0\left(A = \tilde{a} \mid W\right)}{P\left(A = \tilde{a} \mid W\right)} \right] \left[E_P\left(Y \mid A, W\right) - E_0\left(Y \mid A, W\right) \right] \,\middle|\, A = \tilde{a}, W = w \right\} .$$

The above remainder is double robust in the sense that it is zero if either the treatment mechanism (i.e., the probability of A given W) or the outcome regression (i.e., the expected value of Y given A and W) is correctly specified under P. In a randomized trial where the treatment mechanism is known and specified correctly in P, we have that $\mathrm{Rem}_P^R \equiv 0$ and thus $L_P^{RS} \equiv 0$. More generally, an upper bound for Rem_P^R can be found using the Cauchy-Schwarz inequality to relate the rate of $\left\|\mathrm{Rem}_P^R\right\|_{2,P_0}$ to the product of the $L^2(P_0)$-norm for the difference between each of the treatment mechanism and the outcome regression under P and P_0.

Example 2 (Continued).

For (S1) we take $x^R = x^S = w$. Condition (S2) can be verified using an expression similar to that in (4.8). Condition (S3) is satisfied by

$$\begin{aligned} D_P^R(o) &\triangleq \frac{a}{P\left(A = a \mid W = w\right)} \left[y - E_P\left(Y \mid A = a, W = w\right) \right] \\ D_P^S(o) &\triangleq \frac{1-a}{P\left(A = a \mid W = w\right)} \left[y - E_P\left(Y \mid A = a, W = w\right) \right] \ . \end{aligned}$$

If $\min_a P(A = a \mid W)$ is bounded away from zero with probability 1, both $(P, o) \mapsto D_0^R(o)$ and $(P, o) \mapsto D_0^S(o)$ are uniformly bounded.

Similarly to Example 1, we have that $\mathrm{Rem}_P^R(o)$ is equal to

$$E_0 \left\{ \left[1 - \frac{P_0(A = 1 \mid W)}{P(A = 1 \mid W)} \right] [E_P(Y \mid A, W) - E_0(Y \mid A, W)] \,\middle|\, A = 1, W = w \right\}.$$

The remainder $\mathrm{Rem}_P^S(o)$ is equal to the above display but with $A = 1$ replaced by $A = 0$. The discussion about the double robust remainder term from Example 1 applies to these remainders as well.

Example 3 (Continued).

The parameter S is the same as in Example 1. The parameter R satisfies (S1) with $x^R(o) = w$ and (S2) by an identity analogous to that used in Example 1. Condition (S3) is satisfied by $D_P^R(o) \triangleq y - E_P(Y \mid W = w)$. By the bounds on Y, $(P, o) \mapsto D_P^R(o)$ is uniformly bounded. Here, the remainder terms are both exactly zero: $\mathrm{Rem}_P^R \equiv \mathrm{Rem}_P^S \equiv 0$. Thus, we have that $L_P^{RS} \equiv 0$ in this example.

The requirement that $Var_{P_0}\left[D_0^R(O)\right] > 0$ in Corollary 3, and more generally that there exist a nonzero eigenvalue λ_j for the limit distribution in Theorem 14 to be non-degenerate, may at times present an obstacle to our goal of obtaining asymptotic control of the type I error. This is the case for Example 4, which we now discuss further. Nevertheless, we show that with a little finesse the type I error can still be controlled at the desired level for the given test. In fact, the test we discuss has type I error converging to zero, suggesting it may be noticeably conservative in small to moderate samples.

Example 4 (Continued).

In this example, one can take $x^R = w$ and $x^S = w(-k)$. Furthermore, it is easy to show that

$$\begin{aligned} D_P^R(o) &= Y - E_P[Y|W = w] \\ D_P^S(o) &= Y - E_P[Y|W(-k) = w(-k)] \,. \end{aligned}$$

The first-order approximations for R and S are exact in this example as the remainder terms Rem_P^R and Rem_P^S are both zero. However, we note that if $E_P(Y \mid W) = E_P(Y \mid W(-k))$ almost surely, it follows that $D_P^R \equiv D_P^S$. This implies that $\Gamma_0 \equiv 0$ almost surely under $\mathcal{H}_0$. As such, under the conditions of Theorem 14, all of the eigenvalues in the limit distribution of $n\theta_n$ in Theorem 14 are zero and $n\theta_n \to 0$ in probability. We are then no longer able to control the type I error at level α, rendering our proposed test invalid.

Nevertheless, there is a simple albeit unconventional way to repair this example. Let A be a Bernoulli random variable, independent of all other variables, with fixed probability of success $p \in (0, 1)$. Replace S_P with $o \mapsto E_P(Y \mid A = 1, W(-k) = w(-k))$ from Example 2, yielding then

$$D_P^S(o) = \frac{a}{p}\left[y - E_P\left(Y \mid A, W(-k) = w(-k)\right)\right] \ .$$

It then follows that $D_0^R \not\equiv D_0^S$ and in particular Γ_0 is no longer constant. In this case, the limit distribution given in Theorem 14 is non-degenerate. Consistent estimation of $q_{1-\alpha}$ thus yields a test that asymptotically controls type I error. Given that the proposed estimator θ_n converges to zero faster than n^{-1}, the probability of rejecting the null approaches zero as sample size grows. In principle, we could have chosen any positive cutoff given that $n\theta_n \to 0$ in probability, but choosing a more principled cutoff seems judicious.

Because p is known, the remainder term Rem_P^S is equal to zero. Furthermore, in view of the independence between A and all other variables, one can estimate $E_0\left(Y \mid A = 0, W(-k)\right)$ by regressing Y on $W(-k)$ using all of the data without including the covariate A.

4.7 Simulation studies

In simulation studies, we have explored the performance of our proposed test in the context of our Motivating Example, and have also compared our method to the approach of Racine et al. (2006) for which software is readily available – see, e.g., the `R` package `np` (Hayfield and Racine, 2008). We report the results of our simulation studies in this section.

Simulation scenario 1

We use an observed data structure (W, A, Y), where $W \triangleq (W_1, W_2, \ldots, W_5)$ is drawn from a standard 5-dimensional normal distribution, A is drawn according to a Bernoulli(0.5) distribution, and $Y = \mu(A, W) + 5\xi(A, W)$, where the different forms of the conditional mean function $\mu(a, w)$ are given in Table 4.1, and $\xi(a, w)$ is a random variate following a Beta distribution with shape parameters $\alpha = 3\,\text{logit}^{-1}(aw_2)$ and $\beta = 2\,\text{logit}^{-1}[(1-a)w_1]$ shifted to have mean zero.

We performed a test of the null in which $\mu(1, W)$ is equal to $\mu(0, W)$ almost surely as presented in our Motivating Example. Our estimate $\hat{P}_n$ of P_0 was constructed using the knowledge that $P_0\left(A = 1 \mid W\right) = 1/2$, as would be available, for example, in the context of a randomized trial. The conditional mean function $\mu(a, w)$ was estimated using the ensemble learning algorithm Super Learner (van der Laan et al., 2007), as implemented in the `SuperLearner` package (Polley and van der Laan, 2013). This algorithm was implemented using 10-fold cross-validation to determine the best convex combination of regression function candidates minimizing mean-squared error using a candidate library consisting of `SL.rpart`, `SL.glm.interaction`, `SL.glm`, `SL.earth`, and `SL.nnet`. We used the results of Corollary 3 to evaluate significance.

We ran 1000 Monte Carlo simulations with samples of size 125, 250, 500, 1000, and 2000, except for the `np` package, which we only ran for 500 Monte Carlo simulations due to its burdensome computation time. For our Motivating Example we compared our approach

	$\mu(a,w)$	$\overset{a.s.}{=}$
Simulation 1a	$m(a,w)$	$\times$
Simulation 1b	$m(a,w)+0.4[aw_3+(1-a)w_4]$	
Simulation 1c	$m(a,w)+0.8aw_3$	

Table 4.1: Conditional mean function in each of three simulation settings within simulation scenario 1. Here, $m(a,w) \triangleq 0.2\,(w_1^2+w_2-2w_3w_4)$, and the third and fourth columns indicate, respectively, whether $\mu(1,W)$ and $\mu(0,W)$ are equal in distribution or almost surely.

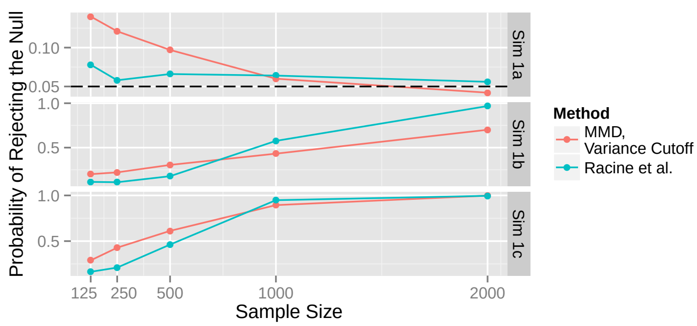

Figure 4.1: Empirical probability of rejecting the null when testing the null hypothesis that $\mu(1,W)-\mu(0,W)$ is equal in probability to zero in Simulation 1.

with that of Racine et al. (2006) using the `npsigtest` function from the `np` package. This requires first selecting a bandwidth, which we did using the `npregbw` function, specifying that we wanted a local linear estimator and the bandwidth to be selected using the `cv.aic` method (Hayfield and Racine, 2008).

Figure 4.1 displays the empirical coverage of our approach as well as that resulting from use of the `np` package. At smaller sample sizes, our method does not appear to control type I error near the nominal level. This is likely because we use an asymptotic result to compute the cutoff, even when the sample size is small. Nevertheless, as sample size grows, the type I error of our test approaches the nominal level. We note that in Racine et al. (2006), unlike in our proposal, the bootstrap was used to evaluate the significance of the proposed test. It will be interesting to see if applying a bootstrap procedure at smaller sample sizes improves our small-sample results. At larger sample sizes, it appears that the method of Racine et al. slightly outperforms our approach in terms of power in simulation scenarios 1a and 1b.

Simulation scenario 2: higher dimensions

We also explored the performance of our method as extended to tackle higher-dimensional hypotheses, as discussed in Section 4.8. To do this, we used the same distribution as for Simulation 1 but with Y now a 20-dimensional random variable. Our objective here

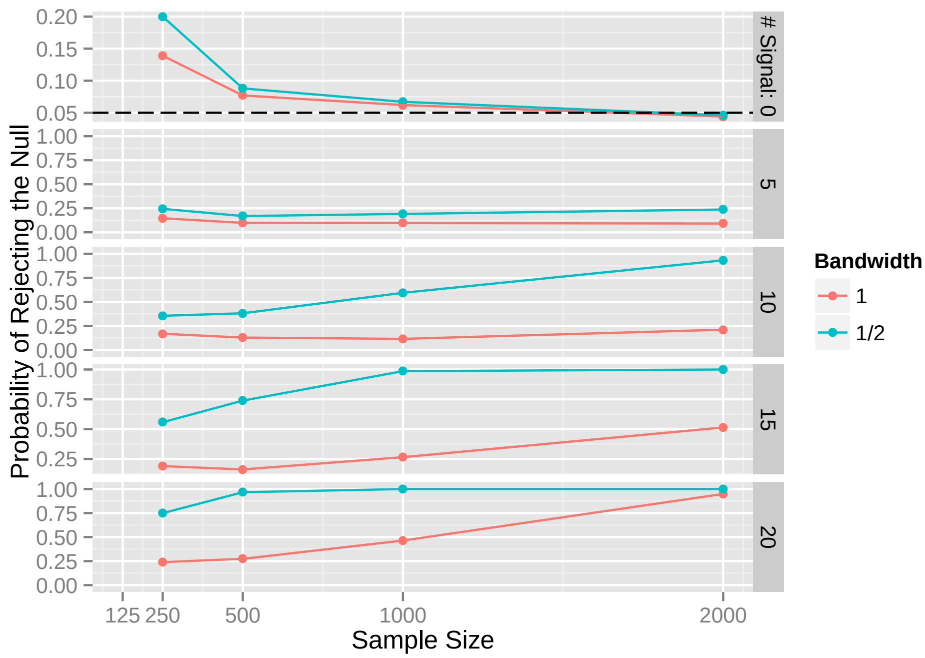

Figure 4.2: Probability of rejecting the null when testing the null hypothesis that $\mu(1, W) - \mu(0, W)$ is equal in probability to zero in Simulation 2.

was to test $\mu(1, W) - \mu(0, W)$ is equal to $(0, 0, \ldots, 0)$ in probability, where $\mu(a, w) \triangleq (\mu_1(a, w), \mu_2(a, w), \ldots, \mu_{20}(a, w))$ with $\mu_j(a, w) \triangleq E_0(Y_j \mid A = a, W = w)$. Conditional on A and W, the coordinates of Y are independent. We varied the number of coordinates that represent signal and noise. For signal coordinate j, given A and W, $20Y_j$ was drawn from the same conditional distribution as Y give A and W in Simulation 1c. For noise coordinate j, given A and W, $20Y_j$ was drawn from the same conditional distribution as Y given A and W in Simulation 1a.

Relative to Simulation 1, we have scaled each coordinate of the outcome to be one twentieth the size of the outcome in Simulation 1. Apart from the Gaussian kernel with bandwidth one, which we have adopted throughout this chapter, we considered defining the MMD with a Gaussian kernel with bandwidth 1/2. Alternatively, this could be viewed as considering bandwidths 1/20 and 1/40 if the outcome had not been scaled by 1/20.

We ran the same Super Learner to estimate $\mu(1, w)$ as in Simulation 1, and we again treated the probability of treatment given covariates as known. We evaluated significance by estimating all of the positive eigenvalues of the centered Gram matrix for $n = 125$ and the largest 200 positive eigenvalues of the centered Gram matrix for $n > 125$.

In Figure 4.2, the empirical null rejection probability is displayed for our proposed MMD method based on bandwidths 1 and 1/2. Our proposal appears to control type I error well at moderate to large sample sizes (i.e., $n \geq 500$). We did not include the results for sample size 125 in the figure because type I error control was too poor. In particular, for zero signal coordinates, the probability of rejection was 0.24 for bandwidth 1 and 0.33 for bandwidth 1/2. For a signal of 5, the empirical probability of rejection decreases between a sample size of 250 and 500, likely due to the poor type I error control at sample size 250. Nonetheless, this simulation shows that, overall, our method indeed has increasing power as sample size grows or as the number of coordinates j for which $\mu_j(1, W) - \mu_j(0, W)$ not equal to zero in probability increases. This figure also highlight that the bandwidth may be an important determinant of finite-sample power, therefore warranting further scrutiny in future work.

4.8 Concluding remarks

We have presented a novel approach to test whether two unknown functions are equal in distribution. Our proposal explicitly allows, and indeed encourages, the use of flexible, data-adaptive techniques for estimating these unknown functions as an intermediate step. Our approach is centered upon the notion of maximum mean discrepancy, as introduced in Gretton et al. (2006), since the MMD provides an elegant means of contrasting the distributions of these two unknown quantities. In their original paper, these authors showed that the MMD, which in their context tests whether two probability distributions are equal using n random draws from each distribution, can be estimated using a U- or V-statistic. Under the null hypothesis, this U- or V-statistic is degenerate and converges to the true parameter value quickly. Under the alternative, it converges at the standard $n^{-1/2}$ rate. Because this parameter is a mean over a product distribution from which the data were observed, it is not surprising that a U- or V-statistic yields a good estimate of the MMD. What is surprising is that we were able to construct an estimator with these same rates even when the null hypothesis involves unknown functions that can only be estimated at slower rates. To accomplish this, we used recent developments from the higher-order pathwise differentiability literature. This appears to be the first use of these developments to address an open methodological problem. Our simulation studies indicate that our asymptotic results are meaningful in finite samples, and that in specific examples for which other methods exist, our methods generally perform at least as well as these established, tailor-made methods. Of course, the great appeal of our proposal is that it applies to a much wider class of problems.

We conclude with several possible extensions of our method that may increase further its applicability and appeal.

1. Although this condition is satisfied in our Motivating Example and two of the three other examples discussed in the introduction, requiring R and S to be in $\mathcal{S}$ can be somewhat restrictive. Nevertheless, it appears that this condition may be weakened by instead requiring membership to $\mathcal{S}^*$, the class of all parameters T for which there exist some $M < \infty$ and elements $T^1, T^2, \ldots, T^M$ in $\mathcal{S}$ such that $T = \sum_{m=1}^{M} T^m$. While the results in our paper can be established in a similar manner for functions in this generalized class, the expressions for the involved gradients are quite a bit more complicated.

Specifically, we find that, for $T, U \in \mathcal{S}^*$ with $T = \sum_{m=1}^{M} T^m$ and $U = \sum_{\ell=1}^{L} U^\ell$, the quantity $\Gamma_P^{TU}(o_1, o_2)$ equals

$$\begin{aligned}
&e^{-[T_P(o_1)-U_P(o_2)]^2} + \sum_{\ell=1}^{L} E_P \left\{ 2\left[T_P(o_1) - U_P(O)\right] e^{-[T_P(o_1)-U_P(O)]^2} \middle| X^{U^\ell} = x_2^{U^\ell} \right\} D_P^{U^\ell}(o_2) \\
&- \sum_{m=1}^{M} E_P \left\{ 2\left[T_P(O) - U_P(o_2)\right] e^{-[T_P(O)-U_P(o_2)]^2} \middle| X^{T^m} = x_1^{T^m} \right\} D_P^{T^m}(o_1) \\
&- \sum_{\ell=1}^{L}\sum_{m=1}^{M} E_{P^2}\Big[\left\{ 4\left[T_P(O_1) - U_P(O_2)\right]^2 - 2 \right\} \\
&\qquad\qquad \times e^{-[T_P(O_1)-U_P(O_2)]^2} \Big| X_1^{T^\ell} = x_1^{T^\ell}, X_2^{U^m} = x_2^{U^m} \Big] D_P^{T^\ell}(o_1) D_P^{U^m}(o_2) \ .
\end{aligned}$$

In particular, we note the need for conditional expectations with respect to X^{R^m} and X^{S^m} in the definition of Γ, which could render the implementation of our method more difficult. While we believe this extension is promising, its practicality remains to be investigated.

2. While our paper focuses on univariate hypotheses, our results can be generalized to higher dimensions. Suppose that $P \mapsto R_P$ and $P \mapsto S_P$ are $\mathbb{R}^d$-valued functions on $\mathcal{O}$. The class $\mathcal{S}_d$ of allowed such parameters can be defined similarly as $\mathcal{S}$, with all original conditions applying componentwise. The MMD for the vector-valued parameters R and S using the Gaussian kernel is given by $\Theta_d(P) \triangleq \Phi_d^{RR}(P) - 2\Phi_d^{RS}(P) + \Phi_d^{SS}(P)$, where for any $T, U \in \mathcal{S}_d$ we set

$$\Phi_d^{TU}(P) \triangleq \iint e^{-\|T_P(o_1)-U_P(o_2)\|^2} dP(o_1) dP(o_2) \ .$$

It is not difficult to show then that, for any $T, U \in \mathcal{S}_d(P_0)$, $\Gamma_{d,P}^{TU}(o_1, o_2)$ is given by

$$\begin{aligned}
&\Big[2\left[T_P(o_1) - U_P(o_2)\right]' \left[D_P^U(o_2) - D_P^T(o_1)\right] + 1 \\
&\quad - 2D_P^T(o_1)' \left\{ 2\left[T_P(o_1) - U_P(o_2)\right]\left[T_P(o_1) - U_P(o_2)\right]' - \mathrm{Id} \right\} D_P^U(o_2) \Big] e^{-\|T_P(o_1)-U_P(o_2)\|^2},
\end{aligned}$$

where Id denotes the d-dimensional identity matrix and A' denotes the transpose of a given vector A. Using these objects, the method and results presented in this chapter can be replicated in higher dimensions rather easily.

3. Our results can be used to develop confidence sets for infinite-dimensional parameters by test inversion. Consider a parameter T satisfying our conditions. Then one can test if $R_0 \triangleq T_0 - f$ is equal in distribution to zero for any fixed function f that does not rely on P. Under the conditions given in this chapter, a $1-\alpha$ confidence set for T_0 is given by all functions f for which we do not reject $\mathcal{H}_0$ at level α. The blip function

from our Motivating Example is a particularly interesting example, since a confidence set for this parameter can be mapped into a confidence set for the sign of the blip function, i.e. the optimal individualized treatment strategy. We would hope that the omnibus nature of the test implies that the confidence set does not contain functions f that are "far away" from T_0, contrary to a test which has no power against certain alternatives. Formalization of this claim is an area of future research.

4. To improve upon our proposal for nonparametrically testing variable importance via the conditional mean function, as discussed in Section 4.6, it may be fruitful to consider the related Hilbert Schmidt independence criterion (Gretton et al., 2005). Higher-order pathwise differentiability may prove useful to estimate and make inferences about this discrepancy measure.

4.9 Proofs

For any $T \in \mathcal{S}$, we will use the shorthand notation $T_t \triangleq T_{P_t}$, $\frac{d}{dt}T_t\big|_{t=\tilde{t}} \triangleq \dot{T}_{\tilde{t}}$ and $\frac{d^2}{dt^2}T_t\Big|_{t=\tilde{t}} \triangleq \ddot{T}_{\tilde{t}}$. Throughout the this section we use the following fluctuation submodel through P_0 for pathwise differentiability proofs:

$$\begin{aligned} dP_t(o) &\triangleq \left(1 + th_1(o) + t^2 h_2(o)\right) dP_0(o), \\ &\text{where } P_0 h_j = 0 \text{ and } \sup_{o\in\mathcal{O}} |h_j(o)| < \infty, j = 1,2. \end{aligned} \tag{4.9}$$

Proofs for Section 4.2

We give two lemmas before proving Theorem 11.

Lemma 6. *For any $T, U \in \mathcal{S}$ and any fluctuation submodel $dP_t = (1 + th_1 + t^2 h_2)\, dP_0$, we have that, for all $\tilde{t}$ in a neighborhood of zero, $\dot{\Phi}^{TU}_{\tilde{t}}$ is equal to*

$$\begin{aligned} &\int \left[\int e^{-[T_{\tilde{t}}(x_1^T) - U_{\tilde{t}}(x_2^U)]^2} dP_{\tilde{t}}(x_1^T)\right] \left[h_1(o_2) + 2\tilde{t}h_2(o_2)\right] dP_0(o_2) \\ &+ \int \left[\int e^{-[T_{\tilde{t}}(x_1^T) - U_{\tilde{t}}(x_2^U)]^2} dP_{\tilde{t}}(x_2^U)\right] \left[h_1(o_1) + 2\tilde{t}h_2(o_1)\right] dP_0(o_1) \\ &- 2\iint \left[T_{\tilde{t}}(x_1^T) - U_{\tilde{t}}(x_2^U)\right] \left[\frac{d}{dt}T_t(x_1^T)\Big|_{t=\tilde{t}} - \frac{d}{dt}U_t(x_2^U)\Big|_{t=\tilde{t}}\right] e^{-[T_{\tilde{t}}(x_1^T) - U_{\tilde{t}}(x_2^U)]^2} dP_{\tilde{t}}(x_2^U) dP_{\tilde{t}}(x_1^T). \end{aligned}$$

Proof of Lemma 6. We have that

$$\begin{aligned}
\dot{\Phi}^{TU}_{\tilde{t}} &= \frac{d}{dt}\iint e^{-[T_t(x_1^T)-U_t(x_2^U)]^2}\left\{\prod_{j=1}^{2}\left[1+th_1(o_j)+t^2h_2(o_j)\right]\right\}dP_0(o_2)dP_0(o_1)\Big|_{t=\tilde{t}} \\
&= \iint \frac{d}{dt}\, e^{-[T_t(x_1^T)-U_t(x_2^U)]^2}\left\{\prod_{j=1}^{2}\left[1+th_1(o_j)+t^2h_2(o_j)\right]\right\}\Big|_{t=\tilde{t}} dP_0(o_2)dP_0(o_1)\ ,
\end{aligned}$$

where the derivative is passed under the integral in view of (S2). The result follows by the chain rule. □

For each $T, U \in \mathcal{S}$, define

$$\begin{aligned}
D^{TU}(o) \triangleq\ & -2\Phi^{TU}(P_0) + \int\left\{2\left[U_0(o_1)-T_0(o)\right]D_0^T(o)+1\right\}e^{-[T_0(o)-U_0(o_1)]^2}dP_0(o_1) \\
& + \int\left\{2\left[T_0(o_1)-U_0(o)\right]D_0^U(o)+1\right\}e^{-[T_0(o_1)-U_0(o)]^2}dP_0(o_1)\ .
\end{aligned}$$

We have omitted the dependence of D^{TU} on P_0 in the notation. We first give a key lemma about the parameter Φ^{TU}.

Lemma 7 (First-order canonical gradient of Φ^{TU})**.** *Let T and U be members of $\mathcal{S}$. Then Φ^{TU} has canonical gradient D^{TU} at P_0.*

Proof of Lemma 7. To consider first-order behavior it suffices to consider fluctuation submodels in which $h_2(o) = 0$ for all o. We first derive the first-order pathwise derivative of the parameter Φ^{TU} at P_0. Applying the preceding lemma at $\tilde{t} = 0$ yields that

$$\begin{aligned}
\frac{d}{dt}\Phi^{TU}(P_t)\Big|_{t=0} &= \int\left[\int e^{-[T_0(x_1^T)-U_0(x_2^U)]^2}dP_0(x_1^T)\right]h_1(o_2)dP_0(o_2) \\
&+ \int\left[\int e^{-[T_0(x_1^T)-U_0(x_2^U)]^2}dP_0(x_2^U)\right]h_1(o_1)dP_0(o_1) \\
&- 2\int\int(T_0(x_1^T)-U_0(x_2^U))(\dot{T}_0(x_1^T)-\dot{U}_0(x_2^U))e^{-[T_0(x_1^T)-U_0(x_2^U)]^2}dP_0(x_2^U)dP_0(x_1^T).
\end{aligned}$$

The first two terms in the last equality are equal to

$$\begin{aligned}
\text{First term} &= \int\left(E_0\left[e^{-[T_0(X^T)-U_0(x^U)]^2}\right]-E_{P_0^2}\left[e^{-[T_0(X_1^T)-U_0(X_2^U)]^2}\right]\right)h_1(o)dP_0(o) \\
\text{Second term} &= \int\left(E_0\left[e^{-[T_0(x^T)-U_0(X^U)]^2}\right]-E_{P_0^2}\left[e^{-[T_0(X_1^T)-U_0(X_2^U)]^2}\right]\right)h_1(o)dP_0(o).
\end{aligned}$$

We now look to find the portion of the canonical gradient given by the third term. We have that

$$\begin{aligned}
-2\int\int(T_0(x_1^T)-U_0(x_2^U))\dot{T}_0(x_1^T)e^{-[T_0(x_1^T)-U_0(x_2^U)]^2}dP_0(x_2^U)dP_0(x_1^T) \\
=\int 2E_0\left[(U_0(X^U)-T_0(x^T))e^{-[T_0(x^T)-U_0(X^U)]^2}\right]D_0^T(o)h_1(o)dP_0(o) \\
2\int\int(T_0(x_1^T)-U_0(x_2^U))\dot{U}_0(x_2^U)e^{-[T_0(x_1^T)-U_0(x_2^U)]^2}dP_0(x_2^U)dP_0(x_1^T) \\
=\int 2E_0\left[(T_0(X^T)-U_0(x^U))e^{-[T_0(X^T)-U_0(x^U)]^2}\right]D_0^U(o)h_1(o)dP_0(o).
\end{aligned}$$

Collecting terms, a first-order Taylor expansion of $t\mapsto\Phi^{TU}(P_t)$ about $t=0$ yields that

$$\Phi^{TU}(P_t)-\Phi^{TU}(P_0)=tE_0\left[D^{TU}(O)h_1(O)\right]+o(t).$$

Thus Φ^{TU} has canonical gradient D^{TU} at P_0. □

The proof of Theorem 11 is simple given the above lemma.

Proof of Theorem 11. Lemma 7, the fact that $\Theta(P)\triangleq\Phi^{RR}(P)-2\Phi^{RS}(P)+\Phi^{SS}(P)$, and the linearity of differentiation immediately yield that the canonical gradient of Θ can be written as $D^{RR}-2D^{RS}+D^{SS}$. Straightforward calculations show that this is equivalent to $o\mapsto 2[P_0\Gamma_0(o,\cdot)-\theta_0]$. □

We will use the following lemma in the proof of Corollary 2 to prove that $R_0(O)$ and $S_0(O)$ are degenerate if $D_1^\Theta\equiv 0$ and $\mathcal{H}_0$ does not hold. Because we were unable to find the proof that the U-statistic kernel for estimating the MMD of two variables X and Y is degenerate if and only if $\mathcal{H}_0$ holds or X and Y are degenerate, we give a proof here that applies in a more general setting than that which we consider in this paper.

Lemma 8. *Let Q be a distribution over $(X,Y)\in\mathcal{Z}^2$, where $\mathcal{Z}$ is a compact metric space. Let $(x,y)\mapsto k(x,y)$ be a universal kernel on this metric space, i.e. a kernel for which the resulting reproducing kernel Hilbert space $\mathcal{H}$ is dense in the set of continuous funtions on $\mathcal{Z}$ with respect to the supremum metric. Further, suppose that $E_Q\sqrt{k(X,X)}$ and $E_Q\sqrt{k(Y,Y)}$ are finite. Finally, suppose that the marginal distribution of X under Q is different from the marginal distribution of Y under Q.*

There exists some fixed constant C such that

$$\int\langle\phi(x_1)-\phi(y_1),\phi(x_2)-\phi(y_2)\rangle_{\mathcal{H}}dQ(x_2,y_2)=C \tag{4.10}$$

for (Q almost) all $(x_1,y_1)\in\mathcal{Z}^2$ if and only if the joint distribution of (X,Y) under Q is degenerate at a single point. Above $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ and $\phi(z)\triangleq k(z,\cdot)$ are the inner product and the feature map in $\mathcal{H}$, respectively.

Proof. If Q is degenerate then clearly (4.10) holds.

If (4.10) holds, then our assumption that X has a different marginal distribution than Y tells us that $C > 0$ (Gretton et al., 2012). Hence, for almost all (x_1, y_1),

$$\langle \phi(x_1) - \phi(y_1), \mu_X - \mu_Y \rangle_{\mathcal{H}} - \langle \mu_X - \mu_Y, \mu_X - \mu_Y \rangle_{\mathcal{H}} = 0,$$

where μ_X and μ_Y in $\mathcal{H}$ have the property that $\langle \mu_X, f \rangle_{\mathcal{H}} = E_Q f(X)$ and $\langle \mu_Y, f \rangle_{\mathcal{H}} = E_Q f(Y)$ for all $f \in \mathcal{H}$ (Lemma 3 in Gretton et al., 2012). The above holds if and only if $\phi(x_1) - \phi(y_1) = \mu_X - \mu_Y$. Noting that $\mu_X - \mu_Y$ does not rely on x_1, y_1, it follows that $\phi(x_1) - \phi(y_1)$ must not rely on x_1, y_1 for all (x_1, y_1) in some Q probability one set $\mathcal{D} \subseteq \mathcal{Z}^2$.

Fix a continuous function $f : \mathcal{Z} \to \mathbb{R}$ and $x_1, y_1 \in \mathcal{D}$. For any $\epsilon > 0$, the universality of $\mathcal{H}$ ensures that there exists an $f_\epsilon \in \mathcal{H}$ such that $\|f_\epsilon - f\|_\infty \leq \epsilon$. By the triangle inequality,

$$|f(x_1) - f(y_1) - f_\epsilon(x_1) + f_\epsilon(y_1)| \leq 2\epsilon.$$

Because $\phi(x_1) - \phi(y_1)$ is constant and $f \in \mathcal{H}$, $\langle \phi(x_1) - \phi(y_1), f_\epsilon \rangle_{\mathcal{H}} = f_\epsilon(x_1) - f_\epsilon(y_1)$ does not rely on x_1, y_1 for any ϵ. Furthermore, the fact that f_ϵ converges to f in supremum norm ensures that $|f_\epsilon(x_1) - f_\epsilon(y_1)|$ converges to a fixed quantity K (which does not rely on x_1 or y_1) as $\epsilon \to 0$. Applying this to the above yields that $f(x_1) - f(y_1) = K$.

As f was an arbitrary continuous function and $X_1 \not\equiv Y_1$, we can apply this relation to $z \mapsto z$ and $z \mapsto z^2$ to show that $x_1 - y_1$ and $x_1 + y_1$ do not rely on the choice of $(x_1, y_1) \in \mathcal{D}$. Hence $(x_1 - y_1 + x_1 + y_1)/2 = x_1$ and $(x_1 + y_1 - y_1 + x_1)/2 = y_1$ do not rely on the choice of $(x_1, y_1) \in \mathcal{D}$. This can only occur if (x_1, y_1) are constant over the probability 1 set $\mathcal{D}$, i.e. if Q is degenerate. □

For the two-sample problem in Gretton et al. (2012), one can take Q to be a product distribution of the marginal distribution of X and the marginal distribution of Y.

Proof of Corollary 2. We first prove sufficiency. If (i) holds, then $2D^{RS} = D^{RR} + D^{SS}$. It follows that $D_1^{\Theta} \equiv 0$ under $\mathcal{H}_0$. Now suppose (ii) holds. It is a simple matter of algebra to verify that $D_1^{RR} \equiv D_1^{RS} \equiv D_1^{SS} \equiv 0$. Hence $D_1^{\Theta} \equiv 0$, yielding the sufficiency of the stated conditions.

We now show the necessity of the stated conditions. Suppose that $\sigma_0 = 0$ and $\mathcal{H}_0$ does not hold. It is easy to verify that

$$\begin{aligned}\tilde{D}_1^{\Theta} \triangleq E_0\left[e^{-[R_0(O)-R_0(o)]^2}\right] + E_0\left[e^{-[S_0(O)-S_0(o)]^2}\right] \\ - E_0\left[e^{-[R_0(O)-S_0(o)]^2}\right] - E_0\left[e^{-[R_0(o)-S_0(O)]^2}\right] - \theta_0\end{aligned}$$

is a first-order gradient in the model where R_0 and S_0 are known (possibly an inefficient gradient depending on the form of R and S). Call the variance of this gradient $\tilde{\sigma}_0$. As the model where R_0 and S_0 are known is a submodel of the (locally) nonparametric model, $\tilde{\sigma}_0 \leq \sigma_0$, and hence $\tilde{\sigma}_0 = 0$ and $\tilde{D}_1^{\Theta} \equiv 0$. Now, if $\tilde{\sigma}_0 = 0$ and $\mathcal{H}_0$ does not hold, then Lemma 8 shows that $R_0(O)$ and $S_0(O)$ are degenerate. Finally, $\tilde{D}_1^{\Theta} \equiv 0$ and the degeneracy of $R_0(O)$ and $S_0(O)$ shows that for almost all o,

$$D_1^{\Theta}(o) = 2D^{RS}(o) = 2(s_0 - r_0)\left(D_0^R(o) - D_0^S(o)\right)e^{-[r_0-s_0]^2},$$

where we use r_0 and s_0 to denote the (probability 1) values of $R_0(O)$ and $S_0(O)$. The above is zero almost surely if and only if $D_0^R \equiv D_0^S$. Thus $\sigma_0 = 0$ only if (i) or (ii) holds. $\square$

We give the following lemma before proving Theorem 12. Before giving the lemma, we define the function $\Pi : \mathcal{S} \to \mathbb{R}$. Suppressing the dependence on P_0 and h_1, h_2, for all $V \in \mathcal{S}$ and $t \neq 0$ we define

$$\begin{aligned}\Pi(V) \triangleq 2\int\int \Big[&2(V_0(o_2) - V_0(o_1))\dot{V}_0(o_2)h_1(o_2) + 2(V_0(o_2) - V_0(o_1))^2\dot{V}_0(o_2)^2 \\ &+ h_2(o_2) - \dot{V}_0(o_2)^2 + (V_0(o_2) - V_0(o_1))\ddot{V}_0(o_2)\Big]e^{-[V_0(o_2)-V_0(o_1)]^2}dP_0(o_2)dP_0(o_1).\end{aligned}$$

Lemma 9. *For any fluctuation submodel consistent with (4.9), $T, U \in \mathcal{S}$ with $T_0(O) \stackrel{d}{=} U_0(O)$, and $t \in \mathbb{R}$ sufficiently close to zero, we have that*

$$\frac{d^2}{dt^2}\Phi^{TU}(P_t)\Big|_{t=0} = 2\int\int \Gamma_0^{TU}(o_1,o_2)h_1(o_1)h_1(o_2)dP_0(o_2)dP_0(o_1) + \Pi(T) + \Pi(U).$$

Proof. Let $H_t(o) \triangleq 1 + th_1(o) + t^2h_2(o)$ and $\dot{H}_t(o) \triangleq h_1(o) + 2th_2(o)$.

$$\begin{aligned}\frac{d^2}{dt^2}\Phi^{TU}(P_t)\Big|_{t=0} = \frac{d}{dt}\int\int \Big[&H_t(o_1)\dot{H}_t(o_2) + \dot{H}_t(o_1)H_t(o_2) \\ &- 2(T_t(o_1) - U_t(o_2))\left(\dot{T}_t(o_1) - \dot{U}_t(o_2)\right)H_t(o_1)H_t(o_2)\Big] \\ &\times e^{-[T_t(o_1)-U_t(o_2)]^2}dP_0(o_2)dP_0(o_1)\Big|_{t=0}\end{aligned} \tag{4.11}$$

We will pass the derivative inside the integral using (S2) and apply the product rule. The first term we need to consider is

$$\begin{aligned}&\frac{d}{dt}\Big[H_t(o_1)\dot{H}_t(o_2) + \dot{H}_t(o_1)H_t(o_2) - 2(T_t(o_1) - U_t(o_2))\left(\dot{T}_t(o_1) - \dot{U}_t(o_2)\right)H_t(o_1)H_t(o_2)\Big]\Big|_{t=0} \\ &= 2\left[h_2(o_1) + h_1(o_1)h_1(o_2) + h_2(o_2)\right] - 2\left(\dot{T}_0(o_1) - \dot{U}_0(o_2)\right)^2 \\ &\qquad - 2(T_0(o_1) - U_0(o_2))\left(\ddot{T}_0(o_1) - \ddot{U}_0(o_2)\right) \\ &\qquad - 2(T_0(o_1) - U_0(o_2))\left(\dot{T}_0(o_1) - \dot{U}_0(o_2)\right)(h_1(o_1) + h_1(o_2)).\end{aligned}$$

The second is

$$\frac{d}{dt}e^{-[T_t(o_1)-U_t(o_2)]^2}\Big|_{t=0} = -2(T_0(o_1) - U_0(o_2))\left(\dot{T}_0(o_1) - \dot{U}_0(o_2)\right)e^{-[T_0(o_1)-U_0(o_2)]^2}.$$

Returning to (4.11), this shows that $\frac{d^2}{dt^2}\Phi^{TU}(P_t)\Big|_{t=0}$ is equal to

$$
\begin{aligned}
&2\int\int\Big[-2(T_0(o_1)-U_0(o_2))\dot{T}_0(o_1)h_1(o_1)+2(T_0(o_1)-U_0(o_2))^2\dot{T}_0(o_1)^2\\
&\qquad+h_2(o_1)-\dot{T}_0(o_1)^2-(T_0(o_1)-U_0(o_2))\ddot{T}_0(o_1)\Big]e^{-[T_0(o_1)-U_0(o_2)]^2}dP_0(o_2)dP_0(o_1)\\
&+2\int\int\Big[2(T_0(o_1)-U_0(o_2))\dot{U}_0(o_2)h_1(o_2)+2(T_0(o_1)-U_0(o_2))^2\dot{U}_0(o_2)^2\\
&\qquad+h_2(o_2)-\dot{U}_0(o_2)^2+(T_0(o_1)-U_0(o_2))\ddot{U}_0(o_2)\Big]e^{-[T_0(o_1)-U_0(o_2)]^2}dP_0(o_2)dP_0(o_1)\\
&+2\int\int\Big[2(T_0(o_1)-U_0(o_2))\Big(\dot{U}_0(o_2)h_1(o_1)-\dot{T}_0(o_1)h_1(o_2)\Big)\\
&\qquad-\big(4(T_0(o_1)-U_0(o_2))^2-2\big)\dot{T}_0(o_1)\dot{U}_0(o_2)+h_1(o_1)h_1(o_2)\Big]e^{-[T_0(o_1)-U_0(o_2)]^2}dP_0(o_2)dP_0(o_1).
\end{aligned}
$$

The expression inside the second pair of integrals only depends on o_1 through $T(o_1)$. Thus we can rewrite this term as $E_0[f(T(O_1))]$ for a fixed function f that relies on P_0, h_1, h_2, and U. Under $\mathcal{H}_0$, we can rewrite this term as $E_0[f(U(O_1))]$. That is, we can replace each $T(O_1)$ in the second pair of integrals with $U(O_1)$. This yields $\Pi(U)$. Switching the roles of o_1 and o_2 in the first pair of integrals above and applying Fubini's theorem shows that

$$
\begin{aligned}
2\int\int&\Big[2(T_0(o_2)-U_0(o_1))\dot{T}_0(o_2)h_1(o_2)+2(T_0(o_2)-U_0(o_1))^2\dot{T}_0(o_2)^2\\
&+h_2(o_2)-\dot{T}_0(o_2)^2+(T_0(o_2)-U_0(o_1))\ddot{T}_0(o_2)\Big]e^{-[T_0(o_2)-U_0(o_1)]^2}dP_0(o_2)dP_0(o_1).
\end{aligned}
$$

By the same arguments used to for the second pair of integrals, the above expression is equal to $\Pi(T)$ under $\mathcal{H}_0$. By (S3), the third pair of integrals can be rewritten as

$$
\begin{aligned}
2\int\int&\Big[2(T_0(o_1)-U_0(o_2))\big(D_0^U(o_2)-D_0^T(o_1)\big)\\
&-\big(4(T_0(o_1)-U_0(o_2))^2-2\big)D_0^T(o_1)D_0^U(o_2)+1\Big]\\
&\times e^{-[T_0(o_1)-U_0(o_2)]^2}h_1(o_1)h_1(o_2)dP_0(o_2)dP_0(o_1).
\end{aligned}
$$

□

Proof of Theorem 12. We start by noting that $\frac{1}{2}\frac{d^2}{dt^2}\theta_t\Big|_{t=0}$ is equal to

$$
\begin{aligned}
&\frac{1}{2}\left[\frac{d^2}{dt^2}\Phi^{TT}(P_t)\Big|_{t=0}+\frac{d^2}{dt^2}\Phi^{UU}(P_t)\Big|_{t=0}-\frac{d^2}{dt^2}\Phi^{TU}(P_t)\Big|_{t=0}-\frac{d^2}{dt^2}\Phi^{UT}(P_t)\Big|_{t=0}\right]\\
&=\int\int\left[\Gamma_0^{RR}(o_1,o_2)+\Gamma_0^{SS}(o_1,o_2)-\Gamma_0^{RS}(o_1,o_2)-\Gamma_0^{SR}(o_1,o_2)\right]h_1(o_1)h_1(o_2)dP_0(o_2)dP_0(o_1)\\
&=\frac{1}{2}\int\int D_2^{\Theta}(o_1,o_2)h_1(o_1)h_1(o_2)dP_0(o_2)dP_0(o_1),
\end{aligned}
$$

where the penultimate equality makes use of Lemma 9. It is easy to verify that $D_2^\Theta(o_1, o_2) = D_2^\Theta(o_2, o_1)$ for all o_1, o_2. The arguments given below the theorem statement in the main text establish the one-degeneracy of Γ_0 under $\mathcal{H}_0$ show that $E_0[D_2^\Theta(O, o)] = E_0[D_2^\Theta(o, O)] = 0$ for all $o \in \mathcal{O}$ under $\mathcal{H}_0$. Condition (S2) ensures that $\left\|D_2^\Theta\right\|_{2,P_0^2} < \infty$, and thus D_2^Θ is P_0^2 square integrable and one-degenerate.

Because the first pathwise derivative is zero under the null, we have that

$$\theta_t - \theta_0 = \frac{1}{2}t^2 \int\int D_2^\Theta(o_1, o_2)h(o_1)h(o_2)dP_0(o_1)dP_0(o_2) + o(t^2).$$

Thus D_2^Θ is a second-order canonical gradient of Θ at P_0. □

We give a lemma before proving Theorem 13.

Lemma 10. *Fix $P \in \mathcal{M}$. For all $T, U \in \mathcal{S}$, let*

$$\mathrm{Rem}_P^{\Phi^{TU}} \triangleq \left\|L_P^{TU}\right\|_{2,P_0} \left\|M_P^{TU}\right\|_{2,P_0} + \left\|\mathrm{Rem}_P^T\right\|_{1,P_0} \left\|\mathrm{Rem}_P^U\right\|_{1,P_0} + \left\|M_P^{TU}\right\|_{4,P_0}^4.$$

There exists a mapping $\zeta(P, P_0, \cdot) : \mathcal{S} \to \mathbb{R}$ such that, for all $T, U \in \mathcal{S}$ for which $T_0(O$ is equal in distribution to $U_0(O)$,

$$\left|P_0^2\Gamma_P^{TU} - \Phi^{TU}(P_0) - \zeta(P, P_0, T) - \zeta(P, P_0, U)\right| \lesssim \mathrm{Rem}_P^{\Phi^{TU}}$$

Proof of Lemma 10. In this proof we use $F(P, P_0, T, U)$ to denote any constant which can be written as $\tilde{\zeta}(P, P_0, T) + \tilde{\zeta}(P, P_0, U)$ for expressions $\tilde{\zeta}(P, P_0, T)$ and $\tilde{\zeta}(P, P_0, U)$ which satisfy $\tilde{\zeta}(P, P_0, T) = \tilde{\zeta}(P, P_0, U)$ whenever $T = U$. We will write

$$c_1F(P, P_0, T, U) + c_2F(P, P_0, T, U) = F(P, P_0, T, U)$$

for any real numbers c_1, c_2. We then fix ζ to be the final instance of $\tilde{\zeta}$ upon exiting the proof.

Fix $T, U \in \mathcal{S}$. Let $b_0(o_1, o_2) \triangleq T_0(o_1) - U_0(o_2)$ and $b(o_1, o_2) \triangleq T_P(o_1) - U_P(o_2)$ for any o_1, o_2. For ease of notation, in the expected values below we will write B and B_0 to refer to $b(O_1, O_2)$ and $b_0(O_1, O_2)$, respectively. We also write T for $T_P(O_1)$, T_0 for $T_0(O_1)$, Rem_P^T for $\mathrm{Rem}_P^T(O_1)$, U for $U_P(O_2)$, U_0 for $U_0(O_2)$, and Rem_P^U for $\mathrm{Rem}_P^U(O_2)$.

We have that

$$\begin{aligned}
P_0^2\Gamma_P^{TU} - \Phi^{TU}(P_0) &= E_{P_0^2}\left[e^{-B^2} - e^{-B_0^2}\right] + E_{P_0^2}\left[2B\left(D_P^U(O_2) - D_P^T(O_1)\right)e^{-B^2}\right] \\
&\quad - E_{P_0^2}\left[\left(4B^2 - 2\right)D_P^T(O_1)D_P^U(O_2)e^{-B^2}\right] \\
&= E_{P_0^2}\left[e^{-B^2} - e^{-B_0^2}\right] - E_{P_0^2}\left[2B\left(B_0 - B\right)e^{-B^2}\right] \\
&\quad + E_{P_0^2}\left[2B\left(\mathrm{Rem}_P^U - \mathrm{Rem}_P^T\right)e^{-B^2}\right] \\
&\quad - E_{P_0^2}\left[\left(4B^2 - 2\right)[T - T_0][U - U_0]e^{-B^2}\right] \\
&\quad - E_{P_0^2}\left[\left(4B^2 - 2\right)\left([T - T_0]\mathrm{Rem}_P^U + \mathrm{Rem}_P^T[U - U_0]\right)e^{-B^2}\right] \\
&\quad - E_{P_0^2}\left[\left(4B^2 - 2\right)\mathrm{Rem}_P^T\,\mathrm{Rem}_P^U\,e^{-B^2}\right].
\end{aligned}$$

A third-order Taylor expansion of $b_0 \mapsto \exp(-b_0^2)$ about $b_0 = b$ yields that $e^{-b^2} - e^{-b_0^2}$ is equal to

$$2b(b_0 - b)e^{-b^2} - \left(2b^2 - 1\right)(b_0 - b)^2 e^{-b^2} + \frac{2}{3} b \left(2b^2 - 3\right)(b_0 - b)^3 e^{-b^2} + O\left((b_0 - b)^4\right),$$

where the magnitude of the $O((b_0 - b)^4)$ term is uniformly bounded above by $C(b_0 - b)^4$ for some constant $C > 0$ when b_0 and b fall in $[-1, 1]$. For the second-order term, we have

$$\begin{aligned} E_{P_0^2}\left[-\left(2B^2 - 1\right)(B_0 - B)^2 e^{-B^2}\right] = {} & E_{P_0^2}\left[\left(4B^2 - 2\right)(T - T_0)(U - U_0)\, e^{-B^2}\right] \\ & - E_{P_0^2}\left[\left([T - T_0]^2 + [U - U_0]^2\right)\left(2B^2 - 1\right) e^{-B^2}\right]. \end{aligned}$$

Thus we have that

$$\begin{aligned} P_0^2 \Gamma_P^{TU} - \Phi^{TU}(P_0) = {} & E_{P_0^2}\left[2B\left(\mathrm{Rem}_P^U - \mathrm{Rem}_P^T\right) e^{-B^2}\right] \\ & + O\left(\|B - B_0\|_{4,P_0}^4\right) - E_{P_0^2}\left[\left(4B^2 - 2\right)\mathrm{Rem}_P^T \,\mathrm{Rem}_P^U\, e^{-B^2}\right] \\ & - E_{P_0^2}\left[\left([T - T_0]^2 + [U - U_0]^2\right)\left(2B^2 - 1\right) e^{-B^2}\right] \\ & + \frac{2}{3} E_0\left[B\left(2B^2 - 3\right)(B_0 - B)^3 e^{-B^2}\right]. \end{aligned} \tag{4.12}$$

A Taylor expansion of $f_1(z) = 2ze^{-z^2}$ shows that there exists a $\tilde{B}_1(o_1, o_2)$ that falls between $B(o_1, o_2)$ and $B_0(o_1, o_2)$ for all o_1, o_2 such that

$$\begin{aligned} & E_{P_0^2}\left[2B\left(\mathrm{Rem}_P^U - \mathrm{Rem}_P^T\right) e^{-B^2}\right] \\ & = E_{P_0^2}\left[\left(\mathrm{Rem}_P^U - \mathrm{Rem}_P^T\right)\left(2B_0 e^{-B_0^2} + (B - B_0)\dot{f}_1(\tilde{B})\right)\right] \\ & = F(P, P_0, T, U) + E_{P_0^2}\left[\left(\mathrm{Rem}_P^U - \mathrm{Rem}_P^T\right)(B - B_0)\dot{f}_1(\tilde{B})\right], \end{aligned} \tag{4.13}$$

where the second equality holds under $\mathcal{H}_0$. The boundedness of $\dot{f}_1$ in $[-2, 2]$, the triangle inequality, and the Cauchy-Schwarz inequality yield

$$\begin{aligned} E_{P_0^2}\left|\left(\mathrm{Rem}_P^U - \mathrm{Rem}_P^T\right)(B - B_0)\dot{f}_1(\tilde{B})\right| & \lesssim E_{P_0^2}\left|\left(\mathrm{Rem}_P^U - \mathrm{Rem}_P^T\right)(B - B_0)\right| \\ & \lesssim E_{P_0^2}\left|L_P^{TU}(O_1) M_P^{TU}(O_2)\right| + E_0\left|L_P^{TU}\right| E_0\left|M_P^{TU}\right| \lesssim \left\|L_P^{TU}\right\|_{2,P_0}\left\|M_P^{TU}\right\|_{2,P_0}. \end{aligned} \tag{4.14}$$

A Taylor expansion of $f_2(z) = (2z^2 - 1)e^{-z^2}$ yields that there exists a $\tilde{B}_2$ that falls between

B and B_0 such that

$$\begin{aligned}
&E_{P_0^2}\left[\left([T-T_0]^2+[U-U_0]^2\right)\left(2B^2-1\right)e^{-B^2}\right] \\
&\quad = E_{P_0^2}\left[\left([T-T_0]^2+[U-U_0]^2\right)\left(2B_0^2-1\right)e^{-B_0^2}\right] \\
&\qquad + 2E_{P_0^2}\left[\left([T-T_0]^2+[U-U_0]^2\right)(B-B_0)\left(B(2B^2-3)\right)e^{-B^2}\right] \\
&\qquad + E_{P_0^2}\left[\left([T-T_0]^2+[U-U_0]^2\right)(B-B_0)^2\frac{\ddot{f}_2(\tilde{B}_2)}{2}\right].
\end{aligned}$$

The first line on the right is equal to $F(P,P_0,T,U)$ under $\mathcal{H}_0$. By the triangle inequality and the boundedness of $\ddot{f}_2$ on $[-2,2]$, the third line satisfies

$$\begin{aligned}
E_{P_0^2}\left[\left([T-T_0]^2+[U-U_0]^2\right)(B-B_0)^2\frac{\ddot{f}_2(\tilde{B}_2)}{2}\right] &\lesssim \sum_{k=0}^{4} E_{P_0^2}\left|[T-T_0]^k\,[U-U_0]^{4-k}\right| \\
\lesssim \sum_{k=0}^{4} E_0\left|[M_P^{TU}]^k\right| E_0\left|[M_P^{TU}]^{4-k}\right| &\lesssim \left\|M_P^{TU}\right\|_{4,P_0}^4 . \qquad (4.15)
\end{aligned}$$

The final inequality above holds by the FKG inequality (Fortuin et al., 1971). It follows that

$$\begin{aligned}
&E_{P_0^2}\left[\left([T-T_0]^2+[U-U_0]^2\right)\left(2B^2-1\right)e^{-B^2}\right]+\frac{2}{3}E_0\left[B\left(2B^2-3\right)(B_0-B)^3e^{-B^2}\right] \\
&\quad =\frac{4}{3}E_{P_0^2}\left[\left([T-T_0]^3-[U-U_0]^3\right)B(2B^2-3)e^{-B^2}\right]+F(P,P_0,T,U)+O(\left\|M_P^{TU}\right\|_{4,P_0}^4) \\
&\quad = F(P,P_0,T,U)+O(\left\|M_P^{TU}\right\|_{4,P_0}^4), \qquad (4.16)
\end{aligned}$$

where the final equality holds under $\mathcal{H}_0$ by a Taylor expansion of $z\mapsto z(2z^2-3)e^{-z^2}$ and analogous calculations to those used in (4.15). We note that the second equality above uses a different F and a different big-O term than the line above, and that the big-O term can be upper bounded by $C\left\|M_P^{TU}\right\|_{4,P_0}^4$ for a constant $C>0$.

Plugging (4.13), (4.14), and (4.16) into (4.12), applying the triangle inequality, and using the bounds on B gives the result. □

We give a lemma before proving Theorem 13.

Lemma 11. *Let $K_P \triangleq \left\|L_P^{RS}\right\|_{1,P_0}+\left\|M_P^{RS}\right\|_{2,P_0}^2$ for all $P\in\mathcal{M}$. If $\mathcal{H}_0$ holds, then for all $P\in\mathcal{M}$,*

$$\sup_{o_1\in\mathcal{O}'}\left|P_0\Gamma_P(o_1,\cdot)\right|\lesssim K_P,$$

where $\mathcal{O}'\subseteq\mathcal{O}$ is some P_0 probability 1 set. More generally, for all $P_0\in\mathcal{M}$,

$$\left|P_0^2\Gamma_P-\theta_0\right|\lesssim K_P.$$

Proof of Lemma 11. For $T, U \in \mathcal{S}$, we have that Γ_P^{TU} is equal to

$$\left[1 + 2(T_P - U_P)D_P^U\right] e^{-[T_P-U_P]^2} - 2\left[(T_P - U_P) + \left(2(T_P - U_P)^2 - 1\right) D_P^U\right] D_P^T e^{-[T_P-U_P]^2}.$$

Above we have omitted the dependence of Γ^{TU} on (o_1, o_2), T and D_P^T on o_1, and U and D_P^U on o_2. For P_0 almost all $o_1 \in \mathcal{O}$, $P_0\Gamma_P^{TU}(o_1, \cdot)$ is equal to

$$\begin{aligned} &P_0\left[1 + 2(T_P(o_1) - U_P)(U_P - U_0)\right] e^{-[T_P(o_1)-U_P]^2} + O\left(\left\|\mathrm{Rem}_P^U\right\|_{1,P_0}\right) \\ &\quad - 2P_0\left[(T_P(o_1) - U_P) + \left(2(T_P(o_1) - U_P)^2 - 1\right)(U_P - U_0)\right] D_P^T(o_1) e^{-[T_P(o_1)-U_P]^2} \end{aligned}$$

where the magnitude of the big-O remainder term is upper bounded by $C\left\|\mathrm{Rem}_P^U\right\|_{1,P_0}$ for a constant $C > 0$ which does not depend on o_1. Taylor expansions of the first and third terms above yield

$$\begin{aligned} P_0\Gamma_P^{TU}(o_1, \cdot) &= P_0 e^{-[T_P(o_1)-U_0]^2} - 2P_0(T_P(o_1) - U_0)D_P^T(o_1)e^{-[T_P(o_1)-U_0]^2} \\ &\quad + O\left(\left\|\mathrm{Rem}_P^U\right\|_{1,P_0}\right) + O\left(\left\|U_P - U_0\right\|_{2,P_0}^2\right), \end{aligned}$$

where the magnitude of the big-O term can be upper bounded by $C\left\|U_P - U_0\right\|_{2,P_0}^2$. If $T_0(O) \overset{d}{=} U_0(O)$, then

$$\begin{aligned} P_0\Gamma_P^{TU}(o_1, \cdot) &= P_0 e^{-[T_P(o_1)-T_0]^2} - 2P_0(T_P(o_1) - T_0)D_P^T(o_1)e^{-[T_P(o_1)-T_0]^2} \\ &\quad + O\left(\left\|\mathrm{Rem}_P^U\right\|_{1,P_0}\right) + O\left(\left\|U_P - U_0\right\|_{2,P_0}^2\right). \end{aligned}$$

Recall that $T, U \in \mathcal{S}$ were arbitrary. Using that $\Gamma_P \triangleq \Gamma_P^{RR} - \Gamma_P^{RS} - \Gamma_P^{SR} + \Gamma_P^{SS}$ and applying the triangle inequality gives the first result.

We now turn to the second result. For any $T, U \in \mathcal{S}$ and $P \in \mathcal{M}$, we have that

$$\begin{aligned} P_0^2\Gamma_P^{TU} =& \Big[2(T_P - U_P)(U_0 - U_P - T_0 + T_P) + 1 \\ &\quad - \left(4(T_P - U_P)^2 - 2\right)(U_P - U_0)(T_P - T_0)\Big] e^{-[T_P-U_P]^2} + O\left(\left\|L_P^{TU}\right\|_{1,P_0}\right) \\ =& \left[2(T_P - U_P)(U_0 - U_P - T_0 + T_P) + 1\right] e^{-[T_P-U_P]^2} \\ &\quad + O\left(\left\|L_P^{TU}\right\|_{1,P_0}\right) + O\left(\left\|M_P^{TU}\right\|_{2,P_0}^2\right) \\ =& \Phi^{TU}(P_0) + O\left(\left\|L_P^{TU}\right\|_{1,P_0}\right) + O\left(\left\|M_P^{TU}\right\|_{2,P_0}^2\right), \end{aligned}$$

where the final equality holds by a first-order Taylor expansion of $(t, u) \mapsto e^{-[t-u]^2}$. The fact that $\Gamma_P \triangleq \Gamma_P^{RR} - 2\Gamma_P^{RS} + \Gamma_P^{SS}$ yields the result. □

Proof of Theorem 13. Fix $P \in \mathcal{M}$ and let P_0 satisfy $\mathcal{H}_0$. We have that

$$\begin{aligned} &P_0^2\Gamma_P - \theta_0 \\ &= P_0^2\Gamma_P^{RR} - \Phi^{RR}(P_0) + P_0^2\Gamma_P^{SS} - \Phi^{SS}(P_0) - \left[P_0^2\Gamma_P^{RS} - \Phi^{RS}(P_0) + P_0^2\Gamma_P^{SR} - \Phi^{SR}(P_0)\right]. \end{aligned}$$

Taking the absolute value of both sides, applying the triangle inequality, and using Lemma 10 yields

$$\begin{aligned}&\left|P_0^2\Gamma_P - \theta_0\right| \\ &\lesssim \operatorname{Rem}_P^{\Phi^{RR}} + \operatorname{Rem}_P^{\Phi^{SS}} + 2\operatorname{Rem}_P^{\Phi^{RS}} \lesssim \left\|L_P^{RS}\right\|_{1,P_0}^2 + \left\|M_P^{RS}\right\|_{4,P_0}^4 + \left\|L_P^{RS}\right\|_{2,P_0}\left\|M_P^{RS}\right\|_{2,P_0},\end{aligned}$$

where the final inequality uses the maximum in the definition of L_P^{RS} and M_P^{RS}.

The inequality for when P_0 satisfies $\mathcal{H}_1$ is proven in Lemma 11. □

Proofs for Section 4.4

Proof of Lemma 5. By the first result of Lemma 11, $|P_0\Gamma_n(o_1,\cdot)| \lesssim K_n$ for P_0 almost all $o_1 \in \mathcal{O}'$. We have that

$$|(P_n - P_0)P_0\Gamma_n| = K_n\left|(P_n - P_0)\left(\frac{P_0\Gamma_n}{K_n}\right)\right|.$$

The fact that $\left\{o_1 \mapsto \frac{P_0\Gamma_n(o_1,\cdot)}{K_n} : \hat{P}_n\right\}$ belongs to a P_0 Donsker class with probability approaching 1 yields that $(P_n - P_0)\left(\frac{P_0\Gamma_n}{K_n}\right) = O_{P_0}(n^{-1/2})$ (van der Vaart and Wellner, 1996), and thus the right-hand side above is $O_{P_0}(K_n/\sqrt{n})$. If $K_n = o_{P_0}(n^{-1/2})$, then this yields that the right-hand side above is $o_{P_0}(n^{-1})$. □

Proof of Theorem 14. Plugging C1), C2), and C3) into (4.4) yields

$$\theta_n - \theta_0 = \mathbb{U}_n\Gamma_0 + o_{P_0}(n^{-1}). \tag{4.17}$$

By Section 5.5.2 of Serfling (1980) and the fact that Γ_0 is P_0 degenerate and uniformly bounded, $n\mathbb{U}_n\Gamma_0 \rightsquigarrow \sum_{k=1}^{\infty}\lambda_k(Z_k^2 - 1)$.

We now prove that all of the eigenvalues of $h(o) \mapsto E_0\left[\tilde{\Gamma}_0(O,o)h(O)\right]$ are nonnegative. Consider a submodel $\{P_t : t\}$ with first-order score $h_1 \in L^2(P_0)$ and second-order score $h_2 \equiv 0$. By the second-order pathwise differentiability of Θ,

$$\frac{\theta_t - \theta_0}{t^2} = \frac{1}{2}\int\int D_2^{\Theta}(o_1,o_2)h_1(o_1)h_1(o_2)dP_0(o_1)dP_0(o_2) + o(1).$$

The left-hand side is nonnegative for all $t \neq 0$ since $\theta_t \geq 0 = \theta_0$ under $\mathcal{H}_0$. Thus taking the limit inferior as $t \to 0$ of both sides shows that

$$\frac{1}{2}\int\int D_2^{\Theta}(o_1,o_2)h_1(o_1)h_1(o_2)dP_0(o_1)dP_0(o_2) \geq 0.$$

Using that $\tilde{\Gamma}_0 = \Gamma_0$ under $\mathcal{H}_0$ and $\Gamma_0 = \frac{1}{2}D_2^{\Theta}$, we have that $\langle o \mapsto E_0[\tilde{\Gamma}_0(O,o)h_1(O)], h_1\rangle \geq 0$, where the inner product is that of $L^2(P_0)$. For any $h_1 \in L^2(P_0)$, it is well known that one can choose a submodel P_t with first-order score $h_1 \in L^2(P_0)$. Hence the above relation holds for all $h_1 \in L^2(P_0)$ and all of the eigenvalues of $h(o) \mapsto E_0\left[\tilde{\Gamma}_0(O,o)h(O)\right]$ are nonnegative. □

Proof of Corollary 3. In this case $\Gamma_0(o_1, o_2) = 2D_0^R(o_1)D_0^R(o_2)$ under $\mathcal{H}_0$. The central limit theorem yields that $\sigma_1^{-1}\sqrt{n}(P_n - P_0)D_0^R \rightsquigarrow Z$. By the continuous mapping theorem,

$$\sigma_1^{-2}n(P_n - P_0)^2\Gamma_0/2 \rightsquigarrow Z^2.$$

Now use that

$$\begin{aligned}\frac{n\mathbb{U}_n\Gamma_0}{2\sigma_1^2} &= \frac{n}{2\sigma_1^2(n-1)}\left[n(P_n - P_0)^2\Gamma_0 - \frac{1}{n}\sum_{i=1}^n \Gamma_0(O_i, O_i)\right] \\ &= \frac{n}{2\sigma_1^2(n-1)}\left[n(P_n - P_0)^2\Gamma_0 - \frac{2}{n}\sum_{i=1}^n D_0^R(O_i)^2\right].\end{aligned}$$

The above quantity converges in distribution to $Z^2 - 1$ by the weak law of large numbers and Slutsky's theorem. □

Proof of Theorem 16. We have

$$\begin{aligned}\theta_n &= 2(P_n - P_0)P_0\Gamma_n + P_0^2\Gamma_n + \mathbb{U}_n\tilde{\Gamma}_n \\ &= 2(P_n - P_0)P_0\Gamma_0 + P_0^2\Gamma_n + \mathbb{U}_n\tilde{\Gamma}_n + 2(P_n - P_0)P_0\left(\Gamma_n - \Gamma_0\right).\end{aligned}$$

By assumption, $\mathbb{U}_n\tilde{\Gamma}_n = o_{P_0}(n^{-1/2})$. The final term is $o_{P_0}(n^{-1/2})$ by the Donsker condition and the consistency condition (van der Vaart and Wellner, 1996). By the second result of Lemma 11 and the assumption that $K_n = o_{P_0}(n^{-1/2})$, this yields that

$$\theta_n - \theta_0 = 2(P_n - P_0)P_0\Gamma_0 + o_{P_0}(n^{-1/2}).$$

Multiplying both sides by $\sqrt{n}$, and applying the central limit theorem yields the result. □

Proof of Corollary 4. We have that

$$\Pr_0\left\{n\theta_n \leq \hat{q}_{1-\alpha}^{ub}\right\} = \Pr_0\left\{\frac{\sqrt{n}(\theta_n - \theta_0)}{\sigma_0} \leq \frac{\hat{q}_{1-\alpha}^{ub}n^{-1/2} - \sqrt{n}\theta_0}{\sigma_0}\right\}$$

Fix $0 < \epsilon < \theta_0$. The right-hand side is equal to

$$\begin{aligned}&\Pr_0\left\{\frac{\sqrt{n}(\theta_n - \theta_0)}{\sigma_0} \leq \frac{\hat{q}_{1-\alpha}^{ub}n^{-1/2} - \sqrt{n}\theta_0}{\sigma_0} \text{ and } \hat{q}_{1-\alpha}^{ub}n^{-1} \leq \epsilon\right\} + o(1) \\ &\leq \Pr_0\left\{\frac{\sqrt{n}(\theta_n - \theta_0)}{\sigma_0} \leq \frac{\sqrt{n}(\epsilon - \theta_0)}{\sigma_0} \text{ and } \hat{q}_{1-\alpha}^{ub}n^{-1} \leq \epsilon\right\} + o(1) \\ &\leq \Pr_0\left\{\frac{\sqrt{n}(\theta_n - \theta_0)}{\sigma_0} \leq \frac{\sqrt{n}(\epsilon - \theta_0)}{\sigma_0}\right\} + o(1) = Pr\left\{Z \leq \frac{\sqrt{n}(\epsilon - \theta_0)}{\sigma_0}\right\} + o(1),\end{aligned}$$

where $Z \sim N(0, 1)$. The final equality holds by Theorem 16 and the well known result about the uniform convergence of distribution functions at continuity points when random variables converge in distribution (see, e.g., Theorem 5.6 in Boos and Stefanski, 2013). The result follows by noting that $(\epsilon - \theta_0)/\sigma_0$ is negative and that $\lim_{z\to-\infty}\Pr(Z \leq z) = 0$. □

Bibliography

J Neyman. On the application of probability theory to agricultural experiments. *Stat. Sci.*, 5:465–480, 1990.

D B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*, 66:688–701, 1974.

J M Robins. A new approach to causal inference in mortality studies with a sustained exposure periodapplication to control of the healthy worker survivor effect. *Math. Model.*, 7(9):1393–1512, 1986.

J Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, New York, 2nd edition, 2009.

S A Murphy. Optimal dynamic treatment regimes. *J. R. Stat. Soc. Ser. B*, 65(2):331–336, 2003.

J M Robins. Discussion of "Optimal dynamic treatment regimes" by Susan A. Murphy. *J. R. Stat. Soc. Ser. B*, 65(2):355–366, 2003.

J M Robins. Optimal structural nested models for optimal sequential decisions. In D Y Lin and Heagerty P, editors, *Proc. Second Seattle Symp. Biostat.*, volume 179, pages 189–326, 2004.

B Chakraborty and E E Moodie. *Statistical Methods for Dynamic Treatment Regimes*. Springer, Berlin Heidelberg New York, 2013.

B Zhang, A A Tsiatis, M Davidian, M Zhang, and E Laber. Estimating optimal treatment regimes from a classification perspective. *Stat*, 68(1):103–114, 2012a.

Y Zhao, D Zeng, A Rush, and M Kosorok. Estimating individual treatment rules using outcome weighted learning. *J. Am. Stat. Assoc.*, 107:1106–1118, 2012.

A R Luedtke and M J van der Laan. Super-learning of an optimal dynamic treatment rule. Technical Report 326, available at `http://www.bepress.com/ucbbiostat/`, Division of Biostatistics, University of California, Berkeley, under review at JCI, 2014a.

J M Robins and A Rotnitzky. Discussion of "Dynamic treatment regimes: Technical challenges and applications". *Electron. J. Stat.*, 8(1):1273–1289, 2014.

M J van der Laan and A R Luedtke. Targeted learning of the mean outcome under an optimal dynamic treatment rule. *J. Causal Inference*, 3(1):61–95, 2014a. doi: 10.1515/jci-2013-0022.

B Zhang, A Tsiatis, M Davidian, M Zhang, and E Laber. A robust method for estimating optimal treatment regimes. *Biometrics*, 68:1010–1018, 2012b.

M J van der Laan and A R Luedtke. Targeted learning of an optimal dynamic treatment, and statistical inference for its mean outcome. Technical Report 329, available at `http://www.bepress.com/ucbbiostat/`, Division of Biostatistics, University of California, Berkeley, 2014b.

B Chakraborty, E B Laber, and Y-Q Zhao. Inference about the expected performance of a data-driven dynamic treatment regime. *Clin. Trials*, 11(4):408–417, 2014.

Y Goldberg, R Song, D Zeng, and M R Kosorok. Comment on Dynamic treatment regimes: Technical challenges and applications. *Electron. J. Stat.*, 8:1290–1300, 2014.

E B Laber, D J Lizotte, M Qian, W E Pelham, and S A Murphy. Rejoinder of Dynamic treatment regimes: Technical challenges and applications. *Electron. J. Stat.*, 8:1312–1321, 2014a.

E Arjas, C Jennison, A P Dawid, D R Cox, S Senn, R G Cowell, V Didelez, Gill R D, J B Kadane, J M Robins, and S A Murphy. Optimal dynamic treatment regimes - Discussion on the paper by Murphy. *J. R. Stat. Soc. Ser. B*, 65:355—-366, 2003.

A Lasry, S L Sansom, K A Hicks, and V Uzunangelov. A model for allocating CDCs HIV prevention resources in the United States. *Health Care Manag. Sci.*, 14(1):115–124, 2011.

G Tao, K Zhao, T Gift, F Qiu, and G Chen. Using a resource allocation model to guide better local sexually transmitted disease control and prevention programs. *Oper. Res. Heal. Care*, 1(2):23–29, 2012.

R M Karp. *Reducibility among combinatorial problems.* Springer, New York Berlin Heidelberg, 1972.

B Korte and J Vygen. *Combinatorial optimization.* Springer, Berlin Heidelberg New York, 5th edition, 2012.

G B Dantzig. Discrete-variable extremum problems. *Oper. Res.*, 5(2):266–288, 1957.

A R Luedtke and M J van der Laan. Optimal dynamic treatments in resource-limited settings. 2015.

A Gretton, M M Borgwardt, M Rasch, B Schölkopf, and A J Smola. A kernel method for the two-sample-problem. In *Adv. Neural Inf. Process. Syst.*, pages 513–520, 2006.

J Pfanzagl. *Asymptotic expansions for general statistical models*, volume 31. Springer-Verlag, 1985.

J M Robins, L Li, E Tchetgen, and A W van der Vaart. Higher order influence functions and minimax estimation of non-linear functionals. In *Essays Honor David A. Free.*, IMS, Collections Probability and Statistics, pages 335–421. Springer New York, 2008.

A W van der Vaart. Higher order tangent spaces and influence functions. *Stat. Sci.*, 29(4): 679–686, 2014.

M Carone, I Díaz, and M J van der Laan. Higher-order Targeted Minimum Loss-based Estimation. 2014.

A R Luedtke and M J van der Laan. Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *Ann. Statist.*, 44(2):713–742, 2016.

P J Bickel, C A J Klaassen, Y Ritov, and J A Wellner. *Efficient and adaptive estimation for semiparametric models.* Johns Hopkins University Press, Baltimore, 1993.

K Hirano and J R Porter. Impossibility results for nondifferentiable functionals. *Econometrica*, 80(4):1769–1790, 2012.

R C Liu and L D Brown. Nonexistence of informative unbiased estimators in singular problems. *Ann. Statist.*, pages 1–13, 1993.

J Chen. Notes on the bias-variance trade-off phenomenon. In *A Festschrift Herman Rubin IMS Lect. Notes Monogr. Ser.*, volume 45, pages 207–217. Institute of Mathematical Statistics, 2004.

E Laber and S Murphy. Adaptive confidence intervals for the test error in classification. *J. Am. Stat. Assoc.*, 106:904–913, 2011.

E B Laber, D J Lizotte, M Qian, W E Pelham, and S A Murphy. Dynamic treatment regimes: Technical challenges and applications. *Electron. J. Stat.*, 8(1):1225, 2014b.

A W van der Vaart and J A Wellner. *Weak convergence and empirical processes.* Springer, Berlin Heidelberg New York, 1996.

K B Athreya and S N Lahiri. *Measure theory and probability theory.* Springer, New York Berlin Heidelberg, 2006.

M J van der Laan and S D Lendle. Online Targeted Learning. Technical Report 330, available at `http://www.bepress.com/ucbbiostat/`, Division of Biostatistics, University of California, Berkeley, 2014.

T Zhang. Solving large scale linear prediction problems using stochastic gradient descent algorithms. In *Proc. 21st Int. Conf. Mach. Learn.*, page 116. ACM, 2004.

J Langford, L Li, and T Zhang. Sparse online learning via truncated gradient. In *Adv. Neural Inf. Process. Syst.*, pages 905–912, 2009.

J Luts, T Broderick, and M P Wand. Real-time semiparametric regression. *J. Comput. Graph. Statist.*, 23(3):589–615, 2014.

P Gaenssler, J Strobel, and W Stute. On central limit theorems for martingale triangular arrays. *Acta Math. Hungar.*, 31(3):205–216, 1978.

J Y Audibert and A B Tsybakov. Fast learning rates for plug-in classifiers. *Ann. Statist.*, 35(2):608–633, 2007.

A B Tsybakov. Optimal aggregation of classifiers in statistical learning. *Ann. Statist.*, 32 (1):135–166, 2004.

R Core Team. *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2014. URL `http://www.r-project.org/`.

W L Steiger. A best possible Kolmogoroff-type inequality for martingales and a characteristic property. *Ann. Math. Stat.*, pages 764–769, 1969.

D A Freedman. On tail probabilities for martingales. *Ann. Probab.*, pages 100–118, 1975.

P Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. ISBN 0-471-19745-9.

A van der Vaart and J A Wellner. Preservation theorems for Glivenko-Cantelli and uniform Glivenko-Cantelli classes. In *High Dimens. Probab. II*, pages 115–133. Springer, 2000.

A R Luedtke and M J van der Laan. Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. Technical Report 332, available at http://biostats.bepress.com/ucbbiostat/paper332/, Division of Biostatistics, University of California, Berkeley, submitted to *Annals of Statistics*, 2014b.

M J van der Laan, E Polley, and A Hubbard. Super Learner. *Stat Appl Genet Mol*, 6(1): Article 25, 2007. ISSN 1.

M J van der Laan and J M Robins. *Unified methods for censored longitudinal data and causality.* Springer, New York Berlin Heidelberg, 2003.

J Pfanzagl. *Estimation in semiparametric models.* Springer, Berlin Heidelberg New York, 1990.

M J van der Laan and S Rose. *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer, New York, New York, 2011.

D W Brock and D Wikler. Ethical challenges in long-term funding for HIV/AIDS. *Health Aff.*, 28(6):1666–1676, 2009.

R Macklin and E Cowan. Given financial constraints, it would be unethical to divert antiretroviral drugs from treatment to prevention. *Health Aff.*, 31(7):1537–1544, 2012.

J A Singh. Antiretroviral resource allocation for HIV prevention. *AIDS*, 27(6):863–865, 2013.

J S Racine, J Hart, and Q Li. Testing the significance of categorical predictor variables in nonparametric regression models. *Econom. Rev.*, 25(4):523–544, 2006.

P Lavergne, S Maistre, and V Patilea. A significance test for covariates in nonparametric regression. *Electron. J. Stat.*, 9:643–678, 2015.

A Berlinet and C Thomas-Agnan. *Reproducing kernel Hilbert spaces in probability and statistics*. Springer Science & Business Media, 2011.

I Steinwart. On the influence of the kernel on the consistency of support vector machines. *J. Mach. Learn. Res.*, 2:67–93, 2002.

J Pfanzagl. *No Title*. Springer, Berlin Heidelberg New York, 1982.

D Nolan and D Pollard. U-processes: rates of convergence. *Ann. Statist.*, 15(2):780–799, 1987.

D Nolan and D Pollard. Functional Limit Theorems for U-Processes. *Ann. Probab.*, 16(3): 1291–1298, 1988. ISSN 0091-1798. doi: 10.1214/aop/1176991691.

A R Luedtke, M Carone, and M J van der Laan. An Omnibus Nonparametric Test of Equality in Distribution for Unknown Functions. *arXiv Prepr. arXiv1510.04195*, 2015.

A Gretton, K Fukumizu, Z Harchaoui, and B K Sriperumbudur. A fast, consistent kernel two-sample test. In *Adv. Neural Inf. Process. Syst.*, pages 673–681, 2009.

C Franz. Discrete approximation of integral operators. *Proc. Am. Math. Soc.*, 134(8):2437–2446, 2006.

G G Gregory. Large sample theory for U-statistics and tests of fit. *Ann. Statist.*, pages 110–123, 1977.

T Hayfield and J S Racine. Nonparametric Econometrics: The np Package. *J. Stat. Softw.*, 27(5), 2008. URL `http://www.jstatsoft.org/v27/i05/`.

E Polley and M J van der Laan. *SuperLearner: super learner prediction*, 2013. URL `http://cran.r-project.org/package=SuperLearner`.

A Gretton, O Bousquet, A Smola, and B Schölkopf. Measuring statistical dependence with Hilbert-Schmidt norms. In *Algorithmic Learn. theory*, pages 63–77. Springer, 2005.

A Gretton, K M Borgwardt, M J Rasch, B Schölkopf, and A Smola. A kernel two-sample test. *J. Mach. Learn. Res.*, 13(1):723–773, 2012.

C M Fortuin, P W Kasteleyn, and J Ginibre. Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.*, 22(2):89–103, 1971.

R J Serfling. *Approximation Theorems of Mathematical Statistics*, volume 37. 1980. ISBN 0471219274. doi: 10.2307/2530199.

D D Boos and L A Stefanski. *Essential Statistical Inference: Theory and Methods*. Springer, Berlin Heidelberg New York, 2013. ISBN 978-1-4614-4817-4.